

**Eine gedächtnispsychologische Analyse
der relativen Effizienz
positiver und negativer Rückmeldung**

INAUGURAL-DISSERTATION
zur
Erlangung des Grades eines Doktors
der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich für Psychologie
der Philipps-Universität Marburg

vorgelegt von
Roland Mix
aus Langenbielau

Marburg/Lahn 1975

**Eine gedächtnispsychologische Analyse
der relativen Effizienz
positiver und negativer Rückmeldung**

INAUGURAL-DISSERTATION

zur

Erlangung des Grades eines Doktors
der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich für Psychologie
der Philipps-Universität Marburg

vorgelegt von
Roland Mix
aus Langenbielau

Marburg/Lahn 1975

Vom Fachbereich für Psychologie der
Philipps-Universität Marburg/Lahn als

Dissertation angenommen am 4.12.1974

Tag der mündlichen Prüfung am 9.1.1975

Berichterstatter: Prof. Dr. K. H. Stapf

Mitberichterstatter: Prof. Dr. Th. Herrmann



Inhaltsverzeichnis

	Seite
1. Einleitung und allgemeine Fragestellung	1
2. Zur "neuen Interpretation" der THORNDIKE-Experimente	8
2.1 Verstärkung vs. Rückmeldung (Terminologie)	8
2.2 Der THORNDIKEsche Standpunkt	10
2.3 Der BUCHWALDSche Standpunkt	16
2.3.1 Zur Bedeutung von "Nichts" (1959)	16
2.3.2 Informationshypothese und Strategieneklärung (1962)	21
2.3.3 Ein probabilistisches Modell zur Wirksamkeit von positiven und negativen Rückmeldungen (1969)	27
2.4 Arbeiten im Anschluß an BUCHWALD (1969)	39
2.5 Latenzzeiten beim BUCHWALD-Modell	45
3. Latenzzeiten: Stand der Forschung	47
3.1 Bereichsabgrenzung	47
3.2 Experimentelle Befunde zu Fehler-Latenzzeiten	51
3.3 Latenzzeitmodelle	57
3.4 Latenzzeiten und Konfidenzurteile	65
4. Latenzzeiten: Hypothesen	67
5. Voruntersuchung: Test des BUCHWALD-Modells	76
5.1 Modifikationen der BUCHWALDschen Versuchsanordnung	76
5.1.1 Beobachtungslernen	76
5.1.2 Weitere Modifikationen	83
5.2 Experiment I: Zum BUCHWALD-Modell	85
5.2.1 Fragestellung und Hypothesen	85
5.2.2 Methode	90
5.2.3 Ergebnisse	101
5.2.4 Diskussion	112

	Seite
5.3 Experiment II: Verschärfte Instruktion und "speed"-Anweisung (k=4)	119
5.3.1 Fragestellung und Hypothesen	120
5.3.2 Methode	121
5.3.3 Ergebnisse	123
5.3.4 Diskussion	125
6. Zwischendiskussion	130
6.1 Folgerungen aus der Voruntersuchung	130
6.2 Kritische Anmerkungen zu BUCHWALDs Arbeiten	131
6.2.1 Notwendigkeit von Nonrehearsal?	132
6.2.2 Einschränkungen der Informationshypothese	134
6.2.3 "Nichts" als neutraler Bezugspunkt?	138
6.3 Neuer Ansatz	141
6.3.1 Modifikation des BUCHWALD-Modells	141
6.3.2 Ein ergänzendes Gedächtnismodell zur Verstärkung	143
6.3.3 Automatisierter Kontrollprozeß und Asymmetrie in Mehrfachwahlaufgaben	145
6.3.4 "Verbalisieren" im Prüfdurchgang	153
7. Hauptuntersuchung	163
7.1 Experiment III: Hauptexperiment	163
7.1.1 Fragestellung und Hypothesen	163
7.1.2 Methode	173
7.1.3 Ergebnisse	188
7.1.4 Diskussion	211
7.2 Ergänzungsuntersuchungen	220
7.2.1 Experiment IV: Striktes Nonrehearsal (k=4)	221
7.2.2 Experiment V: Strategiekontrolle (k=2)	224
7.3 Gesamtauswertung zu den Latenzzeiten (Experimente I bis V)	229
8. Schlußdiskussion	235
8.1 Zu den Latenzzeiten	235
8.2 Zum "Verbalisieren"	238

	Seite
8.3 Zur revidierten Informationshypothese	241
8.4 Kritik	242
9. Zusammenfassung	244
10. Literaturverzeichnis	247
11. Anhang	261
11.1 Lernlisten	261
11.2 Instruktionen	263



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41552302_0007

1. Einleitung und allgemeine Fragestellung

Die vorliegende Arbeit befaßt sich mit der relativen Wirksamkeit von Belohnung und Bestrafung in Humanexperimenten zum verbalen Paarassoziationslernen (PAL). Speziell untersucht sie, welchen Beitrag die Messung von Latenzzeiten zu diesem Problemkreis liefern kann.

Die meisten Arbeiten zur Belohnung und Bestrafung gehen auf die umfangreichen Untersuchungen THORNDIKES (1932) zurück. Ein besonderer Anlaß zu zahlreichen Kontroversen war sein "revidiertes Gesetz des Effekts", in dem eine "asymmetrische" Wirksamkeit positiver und negativer Verstärkungen postuliert wird: Bestrafung soll demnach sehr viel weniger wirksam sein als Belohnung. Einen ausführlichen Überblick über die Nachfolgearbeiten und kritische Analysen geben POSTMAN (1962), BLÖSCHL (1969) und SCHWARZE (1973). Diese Autoren weisen besonders auf die Abhängigkeit der relativen Wirksamkeit von den Randbedingungen, z.B. von der Aufgabenart, hin (s. auch HERRMANN u. STAFF, 1973).

Unter den Diskussionspunkten spielt das sog. Informationsproblem ("Information vs. Effekt") in neuerer Zeit eine besondere Rolle (BLÖSCHL, 1969, S. 58 u. 71; SCHWARZE, 1973, S. 64-94)¹⁾. Die Fragestellung ist: Zu welchem Anteil sind kognitive bzw. motivationale Faktoren im Verstärkungsvorgang beteiligt? Wirken Belohnung und Bestrafung mehr dadurch, daß sie im Lernenden Lust- bzw. Unlustgefühle erzeugen, oder ist es mehr die Information, die in positiven bzw. negativen Verstärkungen²⁾ enthalten ist, die zum Lernfortschritt beiträgt? Der Standpunkt THORNDIKES favorisiert eine Erklärung der Ver-

1) Die Bedeutung des Begriffs "Information" wird in 2.3.2 erklärt. "Information" ist in dieser Arbeit nie im Sinne der Informationstheorie nach SHANNON u. WEAVER (1949) gemeint.

2) Zur genaueren Definition von Verstärkungen und zur Terminologie s. unter 2.1.

stärkungswirkung aufgrund von Lust- bzw. Unlustempfindungen ("satisfying or annoying aftereffects"). Den informativen Standpunkt vertreten u.a. TOLMAN (1932), BUCHWALD (1962, 1969), ESTES (1969), ATKINSON u. WICKENS (1971).

Im folgenden wird besonders die Arbeit von BUCHWALD (1969) betont werden, in der präzise und klar der Informationsstandpunkt in einem mathematischen Modell formuliert wird. Die Absicht der vorliegenden Untersuchung ist also *n i c h t*, festzustellen, *o b* informative Komponenten im Verstärkungsvorgang wirksam sind. Es wird vielmehr eine kognitive Betrachtungsweise vorausgesetzt. Diese wird durch die Struktur der noch zu referierenden Experimente nahegelegt: (a) Es nehmen nur erwachsene Versuchspersonen (Vpn) teil (meistens Studenten). (b) Belohnung und Bestrafung sind die Beurteilungen "richtig" (r) bzw. "falsch" (f) des Versuchsleiters (Vl) zu den Antworten der Vpn. (c) Es gibt nur einen Lerndurchgang, in dem die Vp zunächst die Antworten raten muß; sie kann daher kaum die Rückmeldungen r bzw. f als persönliche Konsequenzen ihrer Leistung, evtl. also emotional auffassen. (Selbstverständlich kann es sich nur um eine Approximation an eine ideal kognitive Versuchssituation handeln.)

Welchen Vorteil hat es nun, von einem mathematisch formulierten Modell zum Verstärkungsvorgang auszugehen? Dieses führt auf die Hauptfragestellung der vorliegenden Untersuchung, nämlich zum möglichen Beitrag von Latenzzeitmessungen zum Problem der relativen Wirksamkeit von Belohnung und Bestrafung. Es ist auffallend, daß in den Literaturübersichten (s.o.) von POSTMAN (1962), BLÖSCHL (1969) und SCHWARZE (1973) keine Arbeiten angeführt werden, in denen die Latenzzeit, also die Zeit zwischen Darbietung des Stimulus und der Reaktion der Vp, eine bedeutsame Rolle spielt. Damit ist zunächst eine theoretische Bedeutsamkeit gemeint, wie sie die Latenzzeitmessung z.B. beim Problem der Alles-oder-Nichts-Hypothese hat (vgl. dazu u.a. WILLIAMS, 1962). Aber auch Latenzzeit als einfaches abhängiges Lernmaß neben anderen Maßen (z.B. FOPPA, 1965, S. 387f.) wird selten angegeben. ESTES (1971, S. 20)

bemerkt dazu, daß bisher Zeitmessungen hauptsächlich in Tierversuchen gemacht wurden, daß er es aber für eine "revolutionäre Chance" hält, Antwortzeiten auch in Humanlernexperimenten zu messen.

Ähnlich äußern sich andere Autoren über die Nützlichkeit und theoretische Bedeutung von Zeitmessungen in Lernexperimenten, z.B. ATKINSON u. SHIFFRIN (1968, S. 160): "Potentially, latencies offer an avenue of analysis that could be more fruitful than the analysis of choice response data..." oder THEIOS (1971, S. 332): "...the most pressing problem facing mathematical models for learning is to account for changes in the shape of the response latency distribution as a function of trials..." oder auch STERNBERG (1969 b, S. 421): "The study of RT should therefore prove helpful to an understanding of the structure of mental activity."

Dieser Mangel an empirischen und theoretischen Arbeiten bezüglich Zeitmessungen in Experimenten zur relativen Wirksamkeit von Belohnung und Bestrafung gab den Anstoß für die vorliegende Untersuchung und folgende Fragestellungen:

- (1) Hängen Latenzzeiten (speziell die Zeitmediane) überhaupt von den Verstärkungsmodalitäten ab, also von r bzw. f ?
- (2) Lassen sich aus einem theoretischen Verstärkungsmodell gerichtete Hypothesen über die mittleren Latenzzeiten ableiten und überprüfen?
- (3) Lassen sich aus den Verhältnissen der Latenzzeitmediane untereinander Aussagen über die relative Effizienz von r und f machen, Aussagen, die eventuell über die "konventioneller" Wiederholungsdaten hinausgehen?

Diese drei Fragen sind nach aufsteigender Bedeutsamkeit geordnet. Die mögliche positive Beantwortung einer Frage hängt jeweils von der positiven Beantwortung der vorhergehenden Frage ab.

Da auf keine empirische und theoretische Basis bezüglich der Verbindung von Latenzzeiten und Verstärkungsproblematik zurückgegriffen werden kann, ist es auf jeden Fall wohl

zweckmäßig, die Zeitmessung in einer experimentellen Situation durchzuführen, die gut ausgearbeitet und hinsichtlich der relativen Wirksamkeit von r und f theoretisch im wesentlichen abgeklärt ist. Diesen Zweck scheinen die Versuchsanordnung und das Modell von BUCHWALD (1969) zu erfüllen. Die Wahl ist auch durch die mathematisch-psychologische Orientierung des Verfassers mitbestimmt. Der Orientierung entspricht eine mögliche Ausweitung bzw. Präzisierung der Frage (2): Falls das BUCHWALD-Modell (BM) durch die Häufigkeitsdaten gut unterstützt werden sollte und die Parameterschätzungen stabil ausfallen, bestünde die Möglichkeit, unter Hinzunahme von einigen Zeitparametern das Modell so zu erweitern, daß genauere Vorhersagen auch bezüglich der Latenzzeitverteilungen zu machen sind, also Vorhersagen, die über eine bloße Rangordnung von Medianen hinausgehen. Eine weitere Fragestellung wäre also:

(2') (a) Ist das BM (1969) hinreichend gut bewährt?

(b) Falls ja: Läßt sich, auf dem BM aufbauend, ein mathematisches Latenzzeitmodell konstruieren?

Wie auch immer (2') beantwortet werden sollte, die Fragestellung (2) bleibt auf alle Fälle für die weitere Untersuchung bestehen.

Falls es sich zeigen sollte, daß das BM als striktes mathematisches Verstärkungsmodell nicht haltbar ist, muß es für die weiteren Betrachtungen modifiziert werden. Die Grundpostulate BUCHWALDs sollten dabei aber möglichst beibehalten werden. Um den Verlust an Exaktheit zu kompensieren, ist dann folgende Strategie zu befolgen: Da die Hauptfragestellung sich auf die Verbindung von Latenzzeiten mit der relativen Wirksamkeit von r und f bezieht, sind Versuchssituationen zu konstruieren, in denen in zuverlässiger und theoretisch abgeklärter Weise "Symmetrie" bzw. "Asymmetrie" in der Wirkung von r und f erzeugt werden kann. Die Latenzzeitmuster unter den verschiedenen symmetrischen bzw. asymmetrischen Bedingungen sind dann miteinander zu vergleichen.

Eine nähere Erläuterung dazu: Im BM wird die relative Effizienz von r und f durch die Axiome erfaßt. Die Vorhersagen

aus einem Latenzzeitmodell nach (2'b) sind daher automatisch relevant bezüglich des Problems: Zeitmessung und relative Effizienz. - Wird (2'a) negativ beantwortet, muß zur Behandlung der Fragen (2) und (3) erst wieder eine sichere Ausgangsbasis bezüglich der relativen Effizienz von r und f geschaffen werden.

Wenn also die Fragestellung (2'b) entfallen muß, sind an ihrer Stelle folgende Fragestellungen zu behandeln:

- (4) Lassen sich in zuverlässiger Weise hinsichtlich der Wirkung von r und f symmetrische bzw. asymmetrische Versuchssituationen herstellen?

Diese allgemeine Fragestellung soll im Rahmen der zur Zeit bevorzugten kognitiven Betrachtungsweise des Verstärkungsvorganges gleich etwas präzisiert werden. Dabei wird auf die Überlegungen BUCHWALDs (1962) Bezug genommen (s. dazu unter 2.3.2):

- (5) (a) Sind für die zuverlässige Herstellung von symmetrischen und asymmetrischen Situationen unterschiedliche Aufgabentypen geeignet, wie z.B. Alternativdiskriminationen, Mehrfachwahlaufgaben, Begriffsbildung?
- (b) Ist die "Informationshypothese", nach z.B. BUCHWALD (1962), ein hinreichend abgeklärtes Konzept zur Erklärung der unterschiedlichen Effizienz von r und f in den obigen Aufgabentypen?¹⁾

In der vorliegenden Arbeit laufen also zwei Problemstellungen parallel:

1. Das Hauptproblem der Abhängigkeit der Latenzzeiten von den Verstärkungsmodalitäten²⁾ r und f (s. Fragen (1) bis (3)).
2. Das Problem der sicheren Ausgangsbasis bezüglich der rela-

1) In vorliegender Untersuchung wird Begriffsbildung zwar mitdiskutiert werden, die Experimente beschränken sich aber aus zeitökonomischen Gründen auf Alternativdiskriminationen und Mehrfachwahlaufgaben.

2) Es wird später auch noch eine dritte ("neutrale") Rückmeldungsmodalität eingeführt werden.

tiven Wirksamkeit von r und f , wie es sich in der Frage (2'a) nach der Gültigkeit des BM (1969) bzw. in den Fragestellungen (4) und (5) ausdrückt.

Da die Frage (2'a) gewissermaßen einen Verzweigungspunkt darstellt, an dem sich entscheidet, ob die Fragestellungen (4) und (5) behandelt werden müssen oder nicht, soll der tatsächliche Ablauf der Arbeit in einem kurzen Überblick dargestellt werden.

Zu Kapitel 2: Da hinsichtlich der relativen Wirksamkeit von Belohnung und Bestrafung der experimentelle und theoretische Rahmen durch die Untersuchungen von THORNDIKE und BUCHWALD aufgespannt wird, werden die Standpunkte der beiden Forscher dargestellt. Hieraus ergeben sich auch die Randbedingungen der vorliegenden Untersuchung. Die Arbeiten von BUCHWALD (1959 a,b; 1962) treten dabei zunächst einmal gegenüber seinem mathematischen Modell von 1969 zurück.

Zu Kapitel 3: Es werden aus der Literatur empirische Befunde zur Latenzzeitmessung und Zeitmodelle dargestellt.

Zu Kapitel 4: Entsprechend der Fragestellung (2) werden gerichtete Hypothesen für die mittleren Latenzzeiten (Mediane) abgeleitet. Die Hypothesen sind nicht an die strikte Gültigkeit des BM (1969) gebunden.

Zu Kapitel 5: In einer Voruntersuchung wird das BM getestet (Fragestellung (2'a)). Um die Rückmeldungshäufigkeiten für r und f gleich zu halten, wird "beobachtendes Lernen" eingeführt. Die experimentellen Daten stützen das Modell nur unzureichend. - Ein Zusatzexperiment ergibt bei den Vpn eine auffällige Schwierigkeit, sich bei Mehrfachwahlaufgaben gleichmäßig auf die Rückmeldungen r bzw. f konzentrieren zu können.

Zu Kapitel 6: Da aufgrund der Voruntersuchung das strikte BM abzulehnen ist, entfällt also die Fragestellung (2'b) zugunsten der Fragestellungen (4) und (5). Im weiteren wird das BM in einer modifizierten Form zugrundegelegt. Die symmetri-

schen bzw. asymmetrischen Versuchssituationen, gemäß den Fragestellungen (4) und (5a), werden mit Hilfe der unabhängigen Variablen (a) Anzahl der Antwortmöglichkeiten ($k=2$ vs. $k=4$) und (b) Lernmodalität ("aktives" vs. "beobachtendes" Lernen) geplant. Zur theoretischen Begründung entsprechend Fragestellung (5b) wird die rein kognitive Informationshypothese partiell in Frage gestellt (Beobachtungslernen nimmt dabei einen höheren Stellenwert ein als in Kapitel 5). Als Prüfmethode wird "Verbalisieren" als eine Kombination von Wiederholungs- und Behaltensprüfung eingeführt.

Zu den Kapiteln 7 und 8: Durchführung des Hauptexperiments. Die Fragestellungen (1), (2), (4) und (5a) werden durch die Ergebnisse positiv beantwortet. Die Beantwortung der Frage (5b) führt zu einer Modifikation der rein kognitiven Informationshypothese. Für Fragestellung (3) ist die positive Beantwortung von (1) und (2) einerseits und (4) und (5a) andererseits Voraussetzung, sie reicht aber nicht für ein bezüglich (3) deutlich günstiges Ergebnis hin.

Abschließend noch eine Bemerkung zu der angegebenen Literatur: Hinsichtlich der allgemeinen Problematik zum Verstärkungsbegriff und insbesondere zur relativen Wirksamkeit von Belohnung und Bestrafung seien zu den drei auf S. 1 angeführten Übersichten noch die Sammelwerke von TAPP (1969) und GLASER (1971) genannt. Darüber hinaus wird in der vorliegenden Arbeit keine auf den letzten Stand gebrachte Vollständigkeit angestrebt. Ausführlicher als in den genannten Werken werden die Arbeiten von BUCHWALD und einschlägige Untersuchungen zum Beobachtungslernen in Paarassoziationsaufgaben behandelt. Bei den referierten Arbeiten zur Latenzzeitproblematik ist ein "bias" zugunsten von mathematisch-psychologischen Untersuchungen zu erkennen, der sich aus der Fragestellung (2'b) leicht erklären läßt.

2. Zur "neuen Interpretation" der THORNDIKE-Experimente

Das Kapitel 2 beinhaltet die Gegenüberstellung der Standpunkte THORNDIKES und BUCHWALDS bezüglich des Verstärkungsproblems. Der kognitive Ansatz BUCHWALDS basiert auf rein gedächtnispsychologischen Annahmen: "In this sense it is a non-reinforcement theory" (BUCHWALD, 1966, S. 85). Die Wahl der obigen Überschrift bezieht sich auf den Titel der BUCHWALDSchen Arbeit von 1969, die hier im Mittelpunkt stehen soll: "Effects of 'Right' and 'Wrong' on subsequent Behavior: A New Interpretation."

2.1 Verstärkung vs. Rückmeldung (Terminologie)

Zunächst sollen einige klärende Bemerkungen zur Terminologie gemacht werden. (Zur genaueren Definitionsfrage s. z.B. BLÖSCHL, 1969, S. 19-28.)

"Es ist im allgemeinen in der Lernpsychologie üblich, ..., unter dem Oberbegriff 'reinforcement' die Begriffspaare 'reward-punishment', 'satisfier-annoyer' und 'positive reinforcer-negative reinforcer' synonym zu verwenden und im Sinne THORNDIKES mit der einen Polarität eine Zunahme der Verhaltenswahrscheinlichkeit, mit der anderen eine Abnahme der Verhaltenswahrscheinlichkeit zu verbinden. In Analogie dazu wird oft für den Belohnungsvorgang der Ausdruck 'positive Verstärkung', für den Bestrafungsvorgang der Ausdruck 'negative Verstärkung' verwendet" (BLÖSCHL, 1969, S. 26f.). An diese Ausdrucksweise soll sich bei der Darstellung des THORNDIKESchen Standpunktes anschließen werden.

Ebenso folgt die vorliegende Untersuchung auch den meisten Arbeiten zur relativen Effizienz von Belohnung und Bestrafung hinsichtlich der Benutzung der Beurteilungen "richtig" (r) und "falsch" (f) seitens des Vl. Man spricht hier von "symbolischen" (auch "verbalen") Verstärkern (POSTMAN, 1962, S. 344). Nach BLÖSCHL (1969, S. 34) gelten die positiven bzw.

negativen Beurteilungen r bzw. f üblicherweise als einander hinsichtlich ihrer Intensität einigermaßen gleich. (Eine Skalierungsstudie von KINTZ u. PAPPAS, 1965, ergibt auf einer "like-dislike"-Dimension für "right" den Wert +2,15 und für "wrong" -1,83.) Soll der informative Aspekt von r bzw. f betont werden, wird von positiver bzw. negativer (Informations-) Rückmeldung gesprochen. BUCHWALD (1969) gebraucht in diesem Zusammenhang den relativ neutralen Ausdruck "outcome".

BERGIUS (1971, S. 73) stellt Begriffe wie Rückkopplung, Kenntnis des Erfolgs, Verstärkung usw. geordnet zusammen: "Zu den Folgen oder Wirkungen des Verhaltens im weitesten Sinne sind recht verschiedene Ereignisse zu zählen:

1. Dynamische Rückkopplung der Bewegung (propriozeptiv, d.h. über die Muskel- und Gelenkrezeptoren).
2. Dynamische Rückkopplung der Bewegung (extrozeptiv, d.h. über die sog. äußeren Sinne, auditiv, visuell, taktil-haptisch: Reafferenz).
3. Statische Rückmeldung = Kenntnis des erzielten Erfolgs (KR, knowledge of results; IF = informatives Feedback).
4. Belohnung, positive Verstärkung (reinforcement).
5. Bestrafung, negative Verstärkung.
6. Weder Belohnung noch Bestrafung.
7. Signal für spätere Verwendbarkeit (NUTTIN u. GREENWALD, 1968)."

Es gibt nur wenige Untersuchungen zum verbalen Lernen, die sich des Rückkopplungsbegriffs im Sinne der Punkte 1. und 2. wenigstens näherungsweise bedienen, also des "feedback"-Begriffs im "eigentlichen" Sinne der Kybernetik. Dazu gehören die Arbeiten von ADAMS (1967), ADAMS u. BRAY (1970), GUTHRIE (1971), WEARING (1971). Dieser "feedback"-Begriff wird im weiteren nicht Verwendung finden, da er eine kontinuierliche Hinführung des IST-Wertes auf den SOLL-Wert beinhaltet, während im Kontext vorliegende Arbeit z.B. die Rückmeldung f eher in einem Alles-oder-Nichts-Sinne wirksam wird.

Die statische Rückmeldung (3.) entspricht der schon erwähnten positiven und negativen (Informations-) Rückmeldung. Belohnung (4.) und Bestrafung (5.) wurden ebenfalls schon erwähnt. Auf Punkt 6 wird unter 2.3.1 einzugehen sein: Dort

soll "weder Belohnung noch Bestrafung" ("Nichts") die Funktion eines neutralen Bezugspunktes erhalten, an dem die relative Wirksamkeit von Belohnung und Bestrafung gemessen werden kann. Der Unterschied zwischen den Bedeutungen nach 3. bzw. 7. ist im Rahmen dieser Untersuchung vernachlässigbar.

Kurz zusammengefaßt: Die Begriffspaare Belohnung/Bestrafung bzw. positive/negative Verstärkung seien für die Darstellung des THORNDIKESchen Standpunktes, positive/negative (Informations-)Rückmeldung für die Darstellung des informationstheoretischen, d.h. z.B. des BUCHWALDSchen Standpunktes, benutzt.

2.2 Der THORNDIKESche Standpunkt

Die folgende Darstellung des THORNDIKESchen Standpunktes lehnt sich an die von BLÖSCHL (1969, S. 16f.) an.

Ausgangspunkt sei das bekannte experimentelle Paradigma eines hungrigen Versuchstiers (Vt; z.B. eine Katze) in einem Problemkäfig (= Reizsituation S), aus dem nur durch einen bestimmten Hebeldruck (= richtige Reaktion R), der die Käfigtür öffnet, zum außerhalb liegenden Futter (= Belohnung) gelangt werden kann. Zu Beginn der Versuchsdurchgänge probiert das Vt alle möglichen Reaktionen seines Verhaltensrepertoires durch, bis es schließlich durch Zufall auf die richtige Reaktion R stößt und damit Erfolg hat, d.h. zum Futter gelangt. Im Verlauf der Versuchsdurchgänge wird dieses "Versuch-und-Irrtums"-Verhalten immer mehr reduziert, bis am Ende das Vt sofort die richtige Reaktion durchführt. Anders ausgedrückt: Die Wahrscheinlichkeit der falschen Reaktionen nimmt immer mehr ab, die der richtigen Reaktion immer mehr zu. Oder noch anders gesagt: Durch den Erfolg (effect) wird die Wiederholungswahrscheinlichkeit der richtigen Reaktion von Durchgang zu Durchgang erhöht, die der falschen Reaktionen infolge des ausbleibenden Erfolges gesenkt.

THORNDIKE interpretierte die Ergebnisse dieser und zahlreicher ähnlicher Untersuchungen u.a. mit Hilfe von folgenden

theoretischen Annahmen:

1. Die Grundlagen des Lernens bestehen in der Verknüpfung (connection) von Reizen (stimuli S; jeweils gesamte Reiz-situation) und Reaktionen (responses R; jeweils komplexe Verhaltenseinheiten), die um so stärker ist, je öfter Reiz und Reaktion gekoppelt auftreten.
2. S-R-Verbindungen, die zu einem für das Individuum angenehmen Resultat führen, werden verstärkt. S-R-Verbindungen, die zu einem für das Individuum unangenehmen Resultat führen, werden abgeschwächt.

Die zweite Annahme bezieht sich also auf die Verhaltenskonsequenzen (im späteren Zusammenhang werden diesedie positiven und negativen "symbolischen" Verstärkungen r und f sein). Bezüglich dieser Verhaltenskonsequenzen stellte THORNDIKE (1911, S. 244) das sog. "theoretische Gesetz des Effekts" auf, in dem eine symmetrische Wirkung von positiver und negativer Verstärkung postuliert wird.

Dem Verstärkungsvorgang liegt nach THORNDIKE eine physiologische Basis zugrunde, sie wird aber nicht näher spezifiziert. THORNDIKE definiert Belohnungen und Bestrafungen ("satisfiers, annoyers") von ihrem Effekt auf das Verhalten her (1911, S. 245):

"By a satisfying state of affairs is meant one which the animal does nothing to avoid, often doing such things as attain and preserve it. By a discomforting or annoying state of affairs is meant one which the animal commonly avoids and abandons."

Nach POSTMAN (1962, S. 344) gehen die "symbolischen" Verstärkungen r bzw. f. konform mit der obigen allgemeinen Definition von Belohnung und Bestrafung:

"Human learners will do nothing to avoid an announcement of 'Right', and often will do such things as to attain and preserve it. And an announcement of 'Wrong' is one which the learner commonly wishes to avoid."

Die Untersuchungen THORNDIKES (1932) führten zu einer Revision des "theoretischen Gesetzes des Effekts" von 1911, und zwar im Sinne einer asymmetrischen Wirkung von positiver und negativer Verstärkung auf S-R-Verbindungen: Bestrafung ist

danach weniger wirksam als Belohnung.

Es sei nun zunächst ein Experiment aus dem Bereich des Humanlernens dargestellt, dessen Ergebnisse typisch für die damaligen Untersuchungen sind und mit anderen ähnlichen Befunden zusammen THORNDIKE veranlaßten, das Postulat der symmetrischen Wirksamkeit aufzugeben (1932, S. 280; Experiment Nr. 72).

Den Vpn wurde an mehreren Tagen hintereinander dieselbe Liste von 200 wenig gebräuchlichen Wörtern (Reizen S) vorgelegt. Jedem Wort dieser Liste waren 5 andere Wörter beigegeben, unter denen die Vpn das Synonym für das Reizwort herausfinden und unterstreichen sollten. Bezüglich eines Reizwortes S stellen also die beigegebenen 5 Wörter das mögliche Reaktionsrepertoire dar. Es gibt in diesem Versuch also zu jedem Reizwort S jeweils $k=5$ Antwortmöglichkeiten (R_i). Nach jeder Wahl sagte der V1 je nach dem Ergebnis "richtig" bzw. "right" oder "falsch" bzw. "wrong". Für zwei Aufgaben (18. und 19.) seien beispielsweise mögliche Abfolgen der Ereignisse (Reizwort, Reaktion, Verstärkung) angegeben:

<u>Reizwort S</u>	<u>Reaktionen R_i</u>	<u>Verstärkung</u>
18. edacious	daring	right
	tractable	
	sober	
	<u>devouring</u>	
	polite	
19. eidolon	-	wrong
	<u>laziness</u>	
	benefice	
	gift	
	duck	
	phantom	

Oder schematischer dargestellt:

$S_{18} - R_4 - \text{right}$ und $S_{19} - R_1 - \text{wrong}$

Zu überprüfen war jeweils im nächsten Durchgang, wie oft als "right" bezeichnete Reaktionen wieder auftraten bzw. als "wrong" bezeichnete Reaktionen nicht mehr auftraten. Nach der ursprünglichen symmetrischen Formulierung des Effektgesetzes (1911) wäre zu erwarten, daß die Wiederholungswahr-

scheinlichkeit richtiger Reaktionen in demselben Maß zunimmt wie die falscher Reaktionen abnimmt. Anders ausgedrückt z.B. für die obigen Aufgaben: "right" stärkt die $S_{18}-R_4$ -Verbindung in demselben Maß, wie "wrong" die $S_{19}-R_1$ -Verbindung abschwächt.

Die Ergebnisse waren nun folgende (z.B. für die Wiederholungen der Reaktionen aus dem zweiten Durchgang im dritten Durchgang): Nach "right" wurden 70% der Reaktionen wiederholt, nach "wrong" 29%. THORNDIKE korrigierte diese Werte noch, indem er eine Zufallswahrscheinlichkeit von 20% (da $k=5$) subtrahierte. "right" bewirkte demnach also einen Anstieg der relativen Wiederholungshäufigkeiten um 50%, "wrong" einen Anstieg um 9%. Bestrafung hatte also eher eine leicht positiv verstärkende als eine unterdrückende Wirkung.

Ähnliche Resultate ergaben sich bei der Berechnung der relativen Wiederholungshäufigkeiten für die übrigen Versuchsdurchgänge des Experiments Nr. 72. Auch in anderen Experimenten innerhalb dieser Versuchsreihe zeigte Belohnung eine deutliche Wirkung, während Bestrafung praktisch wirkungslos blieb (die relativen Wiederholungshäufigkeiten nach "wrong" lagen meistens sogar eher über Zufallsniveau ($= 1/k$), als daß sich ein Unterdrückungseffekt gezeigt hätte).

Die asymmetrischen Ergebnisse waren nicht nur auf Lernexperimente mit verbalem Material beschränkt. In Experiment Nr. 74 z.B. waren zehn motorische Reaktionen (Kopfwendung, Handheben, usw.) zehn inhaltsfreien Strichzeichnungen zuzuordnen.

Aufgrund dieser breiten empirischen Unterstützung formulierte THORNDIKE das "revidierte Gesetz des Effekts" (1932, S. 311): "First, a satisfying after-effect which belongs to a connection can be relied on to strengthen the connection. Second, an annoying after-effect under the same conditions has no such uniform weakening effect...".

Es handelt sich also um ein Gesetz der *a s y m m e t r i s c h e n* Wirksamkeit zugunsten der positiven Verstärkung.

Wegen der Bedeutung für die spätere Diskussion in 6.3.3 sei bezüglich der geringen Wirksamkeit von Bestrafung noch

POSTMAN (1962, S. 349) zitiert:

"When punishment does lead to the modification of behavior, its action is indirect; by inducing variability of behavior it may help to produce the responses which can be strengthened by reward."

Durch Bestrafung einer Reaktion R wird also die zugehörige S-R-Verbindung nicht abgeschwächt, es wird aber eine weitere Reaktion ausgelöst, die nun ihrerseits zu positiver Verstärkung führen kann.

Die referierten Untersuchungen und ihre Interpretation durch THORNDIKE waren der Anlaß für zahlreiche Arbeiten, die sich in versuchsmethodischer und interpretatorischer Hinsicht kritisch mit ihnen auseinandersetzten. Einen Überblick gibt BLÖSCHL (1969, S. 54ff.). Einige Kritikpunkte waren:

1. Problem des neutralen Bezugspunktes: THORNDIKE hatte sich bei der Berechnung der Wirksamkeit von "richtig" (r) und "falsch" (f) auf das Zufallswiederholungsniveau ($= 1/k$) bezogen. Das ist nicht gerechtfertigt, falls V_{pn} z.B. die Tendenz ("response bias") haben, eine einmal gegebene Antwort mit überzufälliger Wahrscheinlichkeit zu wiederholen (z.B. TILTON, 1939), oder falls neutrale Umweltreize an sich schon eine festigende Wirkung auf S-R-Verbindungen haben können (z.B. STEPHENS, 1934a). Beides würde eine größere Wirksamkeit von r vorspiegeln (s. ausführlicher dazu unter 2.3.1).

2. Akzentuierungs- bzw. Isolationsproblem: In THORNDIKES Experimenten (Mehrfachwahlsituationen, also $k > 2$) waren in den ersten Durchgängen die Belohnungen seltener als die Bestrafungen. Es könnte hier ein Isolierungseffekt zugunsten der positiven Verstärkung vorliegen (z.B. ZIRKLE, 1946).

3. Informations- und Strategieneklärung: In einer Mehrfachwahlsituation scheint es lernstrategisch sinnvoll zu sein, sich besonders auf die vom V1 mit r beurteilten Antworten zu konzentrieren, da r eindeutige Information über die richtige Antwort übermittelt, f dagegen nicht (z.B. BUCHWALD, 1962; SPENCE, 1964b; s. ausführlicher unter 2.3.2). SCHWARZE (1973, S. 50f.) schreibt zu diesem Punkt: "Somit scheinen Ergebnisse

aus Mehrfachwahlsituationen wenig geeignet zu sein, das Postulat von der Asymmetrie der Belohnungs- und Bestrafungswirkung empirisch zu bestätigen."

Um den Vergleich mit später zu beschreibenden Versuchsanordnungen zu erleichtern, seien folgende die THORNDIKE-Experimente charakterisierenden Punkte angegeben:

a) Es handelt sich um Mehrfachwahlaufgaben, also $k > 2$.

b) Es wird die "Nichtkorrektur-Methode" (noncorrection procedure) angewandt, d.h. nach einer falschen Reaktion der Vp wird nur "falsch" gesagt, und es wird das nächste Reizwort dargeboten. Die Vp hat also keine Gelegenheit, ihre falsche Reaktion zu korrigieren. POSTMAN (1962, S. 360) spricht hier von einer "vanishing situation", da der Reiz verschwindet, bevor die Vp sich verbessern kann. Er argumentiert:

"According to THORNDIKE, a punishment does not directly weaken the connection which it follows. To test this hypothesis, it is necessary to use situations which minimize the opportunity for the correct response to be performed and reinforced following the punishment. The multiple-choice situations used by THORNDIKE were designed to satisfy this requirement."

Die "Korrektur-Methode" (correction procedure) dagegen, die beim Paarassoziationslernen häufiger angewandt wird, besteht darin, daß jede Reizdarbietung und darauf folgende Reaktion der Vp auf jeden Fall mit der Darbietung der richtigen Antwort beendet wird. Diese Gegenüberstellung von Nichtkorrektur- und Korrektur-Methode (s. z.B. BOWER, 1962) ist wichtig für die Beurteilung von Latenzzeituntersuchungen (s. 3.2). Es sei ferner noch die "Entdeckungs-Methode" (discovery procedure) erwähnt (WITTER et al., 1971), bei der die Vp nach falschen Antworten so lange weiterraten kann, bis sie die richtige Antwort entdeckt hat (s. 6.3.3). Diese Methode wäre also gerade nicht geeignet, entsprechend der obigen Argumentation POSTMANs eine Wirkung von f sichtbar zu machen.

c) Es gibt mehrere Lern- (bzw. Prüf-)Durchgänge. Dem gegenüberzustellen ist das "Ein-Versuchs-Lernen", z.B. bei BUCHWALD (1962, 1969).

d) Nach BLÖSCHL (1969, S. 39) sind die THORNDIKE-Versuche unter "Wiederholungsprüfungen" zu behandeln, da die relevanten Daten Wiederholungshäufigkeiten sind.

e) THORNDIKE legt seinen Berechnungen zur Wirksamkeit von r und f, wie schon oben erwähnt, als Bezugspunkt das Zufallswiederholungsniveau ($1/k$) zugrunde.

2.3 Der BUCHWALDsche Standpunkt

Die Angabe der Erscheinungsjahre der Veröffentlichungen BUCHWALDs in den drei folgenden Überschriften dient im wesentlichen der leichteren Übersicht und der kürzeren Ausdrucksweise bei der späteren Gegenüberstellung der beiden Arbeiten von 1962 und 1969. Unter 2.3.1 bis 2.3.3 werden nicht nur die angegebenen Untersuchungen BUCHWALDs dargestellt, diese geben vielmehr auch Anlaß zu einigen Kommentaren und Literaturangaben, die für die weitere Diskussion wichtig sind. Wie schon in der Einleitung erwähnt, ist BUCHWALD nicht der einzige "Informationstheoretiker". Er wurde aber für die vorliegende Fragestellung ausgewählt, weil sich sein Modell von 1969 besonders gut als Ausgangspunkt für Latenzzeitbetrachtungen eignet. Auf diesem Modell wird zunächst der Schwerpunkt der Diskussionen liegen. Erst in Kapitel 6 wird die Arbeit von 1962 wieder mehr betont. Die beiden Untersuchungen von 1959 (a,b) dienen nur der Einführung der "Nichts"-Rückmeldung.

2.3.1 Zur Bedeutung von "Nichts" (1959)

"Nichts" bezieht sich darauf, daß in den folgenden zu referierenden Experimenten außer den expliziten verbalen Rückmeldungen "richtig" (r) und "falsch" (f) des V1 auch die "Rückmeldung" "Nichts" auftritt. Gemeint ist, daß der V1 nach der Antwort der Vp statt r bzw. f nichts sagt, also keine Stellungnahme über Richtigkeit oder Falschheit der Antwort

abgibt.

Die Begründung für dieses experimentelle Vorgehen leitet sich aus dem unter 2.2 angeführten Problem des neutralen Bezugspunktes her. Die dazu angegebenen Arbeiten demonstrieren u.a., daß schon allein die Tatsache, zu einem Reiz eine Antwort abgegeben zu haben, genügt, beim nächsten Auftreten dieses Reizes die Antwort mit überzufälliger Wahrscheinlichkeit zu wiederholen. Bei der Bestimmung der relativen Wirksamkeit mit Hilfe von Wiederholungshäufigkeiten ergibt dieser Effekt einen Vorteil zugunsten von r.

Der Versuch, diesen Effekt zu kontrollieren, führte zu den sog. "blank-trial"-Experimenten (s. STEPHENS, 1934b, S. 467).

"Unter 'blank-trials' sind dabei nichtverstärkte Reaktionen zu verstehen, d.h. also Antworten der Vp, für die weder Belohnungen noch Bestrafungen erteilt werden. ... Statt der Gegenüberstellung der Effekte von Belohnung und Bestrafung allein handelt es sich hier um den Vergleich der Wirkung von 'Belohnung bzw. Nichts' mit der von 'Bestrafung bzw. Nichts'..." (BLÖSCHL, 1969, S. 83).

Typisch für "blank-trial"-Untersuchungen ist der Vergleich zweier Versuchsgruppen: Die eine Gruppe erhält nach einer richtigen Antwort die explizite Rückmeldung r, nach einer falschen Antwort die Rückmeldung "Nichts" (z.B. Schweigen des V1), kurz: die rb-Gruppe¹⁾. Die andere Gruppe (fb) erhält nach einer richtigen Antwort die Rückmeldung "Nichts" (b), nach einer falschen Antwort die explizite Rückmeldung f. Die Lernerfolge der rb- und der fb-Gruppen sind dann miteinander zu vergleichen. Eventuell wird noch eine dritte Gruppe (rf) zum Vergleich hinzugezogen, die nur explizite Rückmeldungen r und f erhält.

BUSS und Mitarbeiter (1954, 1956, 1959) waren nun aufgrund ihrer Experimente zu folgender Schlußfolgerung gekommen:

1) Zur abkürzenden, von der in der einschlägigen Literatur üblichen leicht abweichenden Schreibweise: Im weiteren werden einheitlich positive Rückmeldungen ("richtig", "right", ...) mit r, negative Rückmeldungen ("falsch", "wrong", ...) mit f, "neutrale" Rückmeldungen ("Nichts", "Nothing", ...) mit b abgekürzt (b von "blank").

"R i g h t is a weaker positive reinforcer than W r o n g is a negative reinforcer, and N o t h i n g is a nonreinforcer" (FERGUSON u. BUSS, 1959, S. 95).

Demnach wäre also der Versuch, einen neutralen Bezugspunkt auf einer Verstärkungsdimension zu finden, gelungen. BUCHWALD (1959 a,b) stellte nun dieser Interpretation seine "acquisition-of-reinforcement-value"-Hypothese entgegen (1959b, S. 355). Sie besagt, daß während des Lernens b in der rb-Kombination negativen, in der fb-Kombination positiven Verstärkungswert gewinnt und dabei im zweiten Fall einen größeren absoluten Verstärkungswert erreicht als im ersten Fall.

Interessant ist hier die Gegenüberstellung der Versuchsanordnungen von 1959a und 1959b. Die Arbeit von 1959a enthält eine typische "blank-trial"-Anordnung, nämlich drei unabhängige Versuchsgruppen unter den Bedingungen rb, fb und rf ("single combination experiment"). In der Untersuchung von 1959b wird in einer einzigen Gruppe jede Vp mit allen drei Verstärkungskombinationen rb, fb und rf konfrontiert ("multiple combination experiment"). Im letzteren Fall dürfte sich nicht, bei Geltung der Hypothese BUCHWALDs, ein einheitlicher Verstärkungswert für b herausbilden, da b ja abwechselnd etwas Positives und Negatives bedeuten würde.

Die Daten ergaben eine starke Stützung der obigen Hypothese.

Die Arbeiten BUCHWALDs (1959 a,b), die hier stellvertretend für viele Arbeiten zum Problem der Bedeutung von b stehen, geben Anlaß zu folgenden Bemerkungen.

Im Bemühen, das Problem des neutralen Bezugspunktes mit Hilfe von "blank-trials" zu lösen, haben sich die Forscher auf diesem Gebiet offensichtlich ein neues Problem geschaffen, eben das der Bedeutung von b. CAIRNS (1967) schreibt dazu: "...it seems reasonable to propose that the experimenter's silence, like other nonverbal events, can function as a positive, negative, or neutral signal, according to its definition within the experimental context" (S. 357).

Ein besonders wichtiger Faktor ist die Art und Weise, in der b in der Instruktion eingeführt wird, ob z.B. die Vp

über die Bedeutung des Schweigens seitens des V1 informiert wird oder nicht. Dazu haben besonders SPENCE und Mitarbeiter (1963, 1964 a u.b, 1965, 1966 a,b,c, 1970, 1972) zahlreiche Beiträge geliefert. Eine sehr sorgfältige methodologische Studie zu diesem Problem stammt von MUELLER u. GUMINA (1972).

Häufig wird das "blank-trial"-Gesetz von LEVINE et al. (1964) angeführt, das salopp ausgedrückt besagt: "Schweigen bedeutet Zustimmung". Danach reagieren Vpn auf b so, als ob der V1 r gesagt hätte. Es scheint allerdings oft nicht beachtet zu werden, daß sich das "blank-trial"-Gesetz besonders auf Problemlösungsaufgaben bezieht, z.B. auf das Beibehalten von Hypothesen nach der "Rückmeldung" b. Nichtsdestoweniger kommt auch SPENCE (z.B. 1972) bei Zugrundelegung anderer Aufgabenarten zu der Folgerung, daß im Falle ungenauer Information Vpn eher die Tendenz haben, b positiv zu interpretieren.

Mit dieser positiven Einschätzung von b seitens der Vpn anscheinend im Widerspruch ist z.B. das Ergebnis von KINTZ u. PAPPAS (1965): Sie erhalten in ihrer Skalierungsstudie für b auf einer "like-dislike"-Dimension den Wert $-1,41$, also einen eindeutig negativen Verstärkungswert. Damit in Übereinstimmung wäre auch AMSELS (1958, 1962) Hypothese der "frustrierenden Nichtbelohnung". Es braucht aber kein echter Widerspruch zu bestehen zwischen den Standpunkten von LEVINE et al. und SPENCE einerseits und den von KINTZ u. PAPPAS und AMSEL andererseits. Bei ersteren sind die Versuchssituationen kognitiv gehalten, es spielen Hypothesen eine Rolle; bei letzteren stehen Emotionen im Vordergrund. Es kann durchaus sein, daß in einer Versuchssituation b von den Vpn als emotional negativ eingeschätzt wird und daß sie trotzdem auf b mit ihren Antworten reagieren wie auf eine positive Rückmeldung.

Eine weitere Bemerkung: Die unterschiedlichen Ergebnisse aus den "single-" bzw. "multiple combination"-Experimenten mahnen zur Vorsicht, Untersuchungen aus dem "blank-trial"-Bereich zu vorschnell miteinander zu vergleichen. In den Arbeiten BUCHWALDS (ab 1962) wird eine weitere Modifikation des üblichen "blank-trial"-Schemas angewandt: Es werden nicht die

Kombinationen rb bzw. fb gegenübergestellt, sondern es werden die Rückmeldungen r, f und b einzeln in ihrer Wirkung auf die Reaktionen der Vpn miteinander verglichen. Die drei Versuchsanordnungen seien noch einmal kurz zusammengestellt:

- 1) Übliches "blank-trial"-Schema: Zwei unabhängige Gruppen, eine unter rb-, die andere unter fb-Bedingungen ("single combination experiment"); evtl. noch eine rf-Gruppe.
- 2) Nur eine Gruppe unter rb-, fb- und rf-Bedingungen ("multiple combination experiment").
- 3) Nur eine Gruppe: Die Rückmeldungen r, f und b treten zufällig gemischt, zu gleichen Anteilen auf (BUCHWALD, ab 1962).

LAHEY u. VERPLANCK (1970) schreiben dazu: "No research has been done to establish the comparability of these various within-and between-S comparisons." Auf den Vergleich von 1) und 2) bezieht sich außer BUCHWALD selbst auch SPENCE (1972).

Besonders problematisch erscheint es, die bisherigen Argumentationen auf die "blank-trial"-Modifikation 3) anzuwenden. Das Problem der unterschiedlichen Bedeutungsverleihung an b im jeweiligen Kontrast zu r bzw. f, wie in den rb- bzw. fb-Gruppen, scheint hier überhaupt wegzufallen. Es könnte sich höchstens um minimale "lokale" Kontrastwirkungen handeln. Es ist hier vor auszuschicken, daß bei BUCHWALD (ab 1962) und in der vorliegenden Arbeit die nicht ganz geklärte k-Abhängigkeit von b wichtiger ist als das in der Literatur bekanntere Problem der Bedeutungskontrastierung von b gegenüber r bzw. f.

Was die Hypothese der präexperimentellen eher positiven Bedeutung von b betrifft (s. z.B. das "blank-trial"-Gesetz), so findet sich auch in Untersuchungen entsprechend Modifikation 3) eine überzufällige Wiederholungstendenz nach b. Diese wird im Zusammenhang mit dem "g-Mechanismus" in 2.3.3 diskutiert werden.

Abschließend noch eine Bemerkung zu der experimentellen

Darstellung von b: Bei den meisten einschlägigen Untersuchungen schweigt der Vp, wenn b als "Rückmeldung" erforderlich ist. Das muß aber nicht unbedingt sein. Vielmehr scheint eine Variation in der experimentellen Darstellung von b keinen bedeutsamen Einfluß zu haben, falls nur die Instruktion hinreichend klar ist (z.B. SPENCE, 1972).

STEPHENS (1934b) bot seinen Vpn in einer Mehrfachwahlsituation die Rückmeldungen visuell an (nach BLÖSCHL, 1969, S. 65): b entweder durch ein leeres Blatt oder durch ein "irrtümlich" getipptes Zeichen. Die Wiederholungshäufigkeiten nach diesen beiden Formen von b waren 56% bzw. 58%. (Die Differenz ist sehr gering gegenüber den Wiederholungshäufigkeiten nach r bzw. f: 77% bzw. 44%)

RIMM et al. (1972) verglichen in einer Alternativwahlsituation die beiden Darstellungen von b "Schweigen" und "redundante Erinnerung an die beiden möglichen Antwortalternativen" miteinander. Es ergab sich kein signifikanter Unterschied. Auf diese Ergebnisse wird unter 5.1.2 Bezug genommen werden.

2.3.2 Informationshypothese und Strategieneklärung (1962)

Das Hauptanliegen BUCHWALDS in seinen beiden Experimenten von 1962 war "...to determine an empirical baseline of response repetition in the absence of both 'reward' and 'punishment'...". Er wählte dazu die in 2.3.1 angegebene Modifikation 3) des "blank-trial"-Schemas. Es ergab sich, daß die beiden von ihm benutzten Methoden zur Bestimmung der empirischen Bezugslinie ("PIT"- und "IT"-Prozedur) sich nicht signifikant unterschieden. Der Einfachheit halber soll daher nur die "IT"-Prozedur beschrieben werden. "IT" bedeutet: Es wird nur ein Lerndurchgang (information trial I) und ein Prüfdurchgang (test trial T) durchgeführt.

Experiment Nr. 1: Die Versuchsanordnung ähnelte der in den THORNDIKESchen Arbeiten. Es handelte sich um Mehrfachwahlaufgaben ($k=4$). Den Vpn wurde eine Liste von 40 Wörtern vorgelesen, auf die sie jeweils mit einer der Zahlen 1, 3, 5 und 7

zu reagieren hatten. Der Wichtigkeit halber sollen BUCHWALDS Anmerkungen zur Instruktion wörtlich zitiert werden (1962, S. 72):

"The S was told that he would be read a list of words, that one of the numbers 1, 3, 5, 7 had been randomly assigned to each word, that his task was to learn which number went with which word, and that he would be given a test trial to show what he had learned. The S was instructed to say one of the four numbers each time E said a word. He was also told that E would indicate a correct (incorrect) response by saying Right (Wrong) and that when E said nothing it meant that S was being given no information." (Unterstreichungen vom Verf.)

Die letzte Unterstreichung betont nur, daß die Bedeutung von b durch diese Instruktion klargestellt wird. Auf die erste Unterstreichung wird noch in 6.2.2 zurückzukommen sein.

In der Liste von 40 Wörtern waren die beiden ersten und letzten nur "buffer words", die übrigen 36 waren von vornherein, ohne Wissen der Vp natürlich, im Verhältnis 12:12:12 als r-, f- und b-Wörter festgelegt und zufällig gemischt. Dadurch sollte ein möglicher Isolierungseffekt ausgeschaltet werden (z.B. ZIRKLE, 1946; POSTMAN u. ADAMS, 1955). Im auf den Lern-durchgang folgenden Prüfdurchgang gab der V1 keine Rückmeldung mehr.

Verglichen wurden die relativen Häufigkeiten, mit denen die Antworten auf r-, f- und b-Wörter im Prüfdurchgang jeweils wiederholt wurden. Die relativen Wiederholungshäufigkeiten (in %) betrugen (nach BLÖSCHL, 1969, S. 98) rr = 41% nach r, bb = 33% nach b und ebenso auch ff = 33% nach f.

Im weiteren soll diese abkürzende Schreibweise: rr, bb und ff, für die relativen Wiederholungshäufigkeiten nach r, b und f beibehalten werden. Nichtwiederholungen werden entsprechend durch r $\bar{r}$, b $\bar{b}$ und f $\bar{f}$ zu kennzeichnen sein. Wo keine Verwechslungen zu befürchten sind, werden diese sechs Symbole auch als Indices verwandt, z.B. bei den Latenzzeiten oder als allgemeine Hinweise auf das (Nicht-)Wiederholen von Reaktionen nach r, b und f.

Die oben zitierten Ergebnisse seien in einem Diagramm (s. Abb. 2.3.2-1) dargestellt, wie es in der Form noch häufig

auftreten wird. Auf der Ordinatenachse sind die relativen Wiederholungshäufigkeiten (in %) nach r, b und f, also rr, bb und ff, aufgetragen. Die Abstände auf der Abzissenachse haben zwar keine Bedeutung, die Rangfolge ist aber insoweit wichtig, als b, der Absicht nach, auf einem hypothetischen Verstärkungscontinuum zwischen r und f liegen sollte.

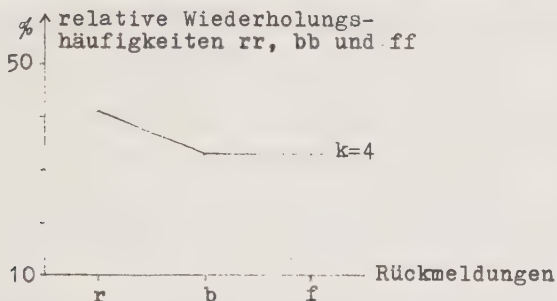


Abb. 2.3.2-1: Veranschaulichung der Ergebnisse BUCHWALDS (1962, Experiment Nr. 1)

Die Wirksamkeit von r bzw. f müßte sich daran bemerkbar machen, daß rr möglichst höher und ff möglichst niedriger als das bb-Niveau sein sollten. Ergibt sich kein monotoner Abfall des "rr-bb-ff-Profiles", liegt entweder eine starke Zufallsschwankung vor, oder die Einführung von b hat sich als Fehlschlag erwiesen.

Das obige Ergebnis zeigt nun, daß die Wirksamkeit von f, bezogen auf das bb-Niveau, praktisch gleich Null war. In früheren Arbeiten nach dem "blank-trial"-Schema (z.B. BUSS et al., 1956) hatte sich die fb-Kombination wirksamer als die rb-Kombination erwiesen. Wie kommt es nun zu diesem diskrepanten Ergebnis bezüglich der Wirksamkeit von f? BUCHWALD (1962) argumentiert vom Informationsgehalt her, der von den Rückmeldungen r bzw. f übermittelt wird und der von der Aufgabenart abhängt.

Bei den Arbeiten von BUSS et al. (1956) handelt es sich um Begriffsbildungsaufgaben: Dort wirken r bzw. f durch Bestätigung oder Verwerfung von Hypothesen. Ein r seitens des V1

kann aber erfolgen, obwohl die Vp noch aufgrund einer falschen Hypothese antwortet; ein f hingegen gibt eindeutige Information über die Falschheit der augenblicklichen Hypothese: Diese ist also zu verwerfen und eine neue zu probieren. Hier ist demnach f wirksamer als r (s. dazu auch MEYER u. OFFENBACH, 1962; LEVINE et al., 1964; JONES, 1961).

Bei Mehrfachwahlaufgaben (z.B. $k=4$) ist es gerade umgekehrt: Hier gibt r eindeutige Information, während f keinen direkten Hinweis auf die richtige Antwort gibt (s. dazu auch TILTON, 1939; ZIRKLE, 1946).

Sollte diese "Informationshypothese" begründet sein, müßten in einer Alternativwahlsituation ($k=2$), in der ja r und f gleiche Information bezüglich der richtigen Antwort liefern, r und f gleich wirksam sein.

Zur Überprüfung dieser Vorhersage führte BUCHWALD das Experiment Nr. 2 durch, in dem sowohl $k=4$ - als auch $k=2$ -Aufgaben zu lernen waren. Als Wiederholungshäufigkeiten ergaben sich für $k=2$: rr = 54%, bb = 56%, ff = 49%; für $k=4$: rr = 35%, bb = 31%, ff = 28%.

(Die zu diesen Ergebnissen gehörenden "Profile" sind in Abb. 2.3.2-2 dargestellt.)

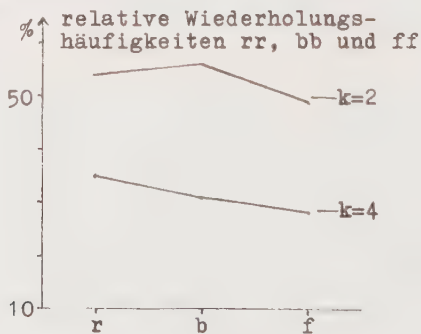


Abb. 2.3.2-2: Veranschaulichung der Ergebnisse BUCHWALDS (1962, Experiment Nr. 2)

Die Resultate für $k=4$ entsprechen im wesentlichen denen des Experimentes Nr. 1. Daß bei $k=2$ bb höher liegt als rr, ist

als Zufallsschwankung zu betrachten: Profile aus späteren Experimenten zeigen einen "regulären" Abfall von rr über bb nach ff. Auf jeden Fall liegt bei $k=2$ der ff-Wert signifikant niedriger als die rr- und bb-Werte, die sich nicht bedeutsam unterscheiden. Bezieht man die relative Effizienz von r und f wieder auf das bb-Niveau, so ist bei $k=2$ r wirkungslos, während f einen markanten Unterdrückungseffekt zeigt. Die Verhältnisse sind demnach umgekehrt wie bei $k=4$.

Diese totale Umkehrung ist nicht mit der obigen Informationshypothese vereinbar, nach der bei $k=2$ r und f gleich wirksam sein müßten. Zur Erklärung dieser Diskrepanz führt BUCHWALD zusätzlich zu der Informationshypothese noch evtl. von k abhängige Strategien der Vpn an: Bei $k>2$ -Aufgaben konzentrieren sich die Vpn im Lerndurchgang auf S-R-Verbindungen, die von r gefolgt wurden, und vernachlässigen S-R-f-Sequenzen. Bei $k=2$ -Aufgaben konzentrieren sich die Vpn besonders auf die S-R-Verbindungen, die von f gefolgt wurden, um sich im Prüfdurchgang verbessern zu können. Die Möglichkeit des effektiven Verbesserns besteht ja in der $k=2$ -Situation.

Durch diese Zusatzhypothese der unterschiedlichen Strategien lassen sich tatsächlich die obigen Ergebnisse erklären. BLÖSCHL (1969, S. 164) findet diese Strategieneklärung bei $k=2$ allerdings "wenig plausibel". (BUCHWALD selbst hat zu diesem Problem keine weiteren Experimente durchgeführt.) Auf jeden Fall ist bei der Interpretation dieser Resultate sauber zwischen der Informationshypothese und der zusätzlichen Strategieneklärung zu unterscheiden!

Als eine Alternativerklärung für die von k abhängige unterschiedliche Effizienz von r und f führt BUCHWALD noch eine mögliche Abhängigkeit des bb-Wiederholungsniveaus von k an. Es fällt zunächst auf, daß die bb-Werte deutlich über dem jeweiligen Zufallsniveau liegen: Bei $k=2$ ist $bb = 56\% > 50\% = 1/2$, und bei $k=4$ ist $bb = 31\% > 25\% = 1/4$. Es scheinen gewisse systematische Faktoren (z.B. Gedächtniseinflüsse oder präexperimentelle S-R-Verbindungen) im Spiel zu sein, die eine überzufällige Wiederholung nach b bewirken. Wenn diese systemati-

schen Faktoren im Falle $k=2$ stärker wirksam sein sollten als bei $k>2$, wären die unterschiedlichen Wirksamkeiten von r und f als Funktion von k auch plausibel gemacht.

In einer kleinen Rechnung versucht BUCHWALD diese mögliche Erklärung zu stützen. Der sich ergebende Effekt ist aber nur klein und ändert nichts an der Asymmetrie von r und f zugunsten von f im Falle $k=2$. Ungeklärt bleibt auch, warum gerade bei $k=2$ die systematischen Faktoren wirksamer sein sollten als bei $k>2$. (U.a. wegen dieser unsicheren Ausgangsbasis bezüglich der k -Abhängigkeit von bb wird später ein Maß für die relative Effizienz von r und f vorgeschlagen werden, das nicht mehr von dem bb -Niveau abhängt.)

Abschließend ist zusammenzufassen:

1. Die Ergebnisse für die Mehrfachwahlaufgaben ähneln denen THORNDIKES (d.h. Wirkungslosigkeit von f), BUCHWALD interpretiert sie aber durch kognitive Faktoren wie Informationsverarbeitung und unterschiedliche Strategien.
2. Der Ort der Entstehung der unterschiedlichen Effizienz von r bzw. f wird auf das strategienabhängige Einspeichern ("storing") des Lernmaterials im Lerndurchgang verlegt. - Im nächsten Abschnitt (2.3.3) entsteht die Asymmetrie erst im Prüf-durchgang, also während des Wiederfindens des Gelernten ("retrieval") und bei der Entscheidung zu antworten.
3. In der Arbeit von 1962 wird die relative Effizienz von r und f nur in bezug auf das empirische bb -Niveau beurteilt und diskutiert. BUCHWALD (1969) betont dagegen mehr die Beziehungen zum Zufallsniveau ($1/k$).
4. Im Prinzip stehen Informationshypothese und Strategieneklärung im Widerspruch mit den Postulaten des im nächsten Abschnitt darzustellenden Modells von BUCHWALD (1969). Dieser Widerspruch stellt aber keine Schwierigkeit dar, wenn es gelingt, "rehearsal" im Anschluß an die Rückmeldungen zu verhindern. Dann können nämlich unterschiedliche Strategien bzw. Aufmerksamkeitszuwendungen nicht wirksam werden (zur weiteren Diskussion s. unter 6.2.2). Unter diesem Vorbehalt wird jetzt vorläufig ausschließlich das Modell (1969) behandelt werden.

3.3 Ein probabilistisches Modell zur Wirksamkeit von positiven und negativen Rückmeldungen (1969)

In seiner Arbeit von 1969 faßt BUCHWALD die Ergebnisse und Überlegungen aus seinen Untersuchungen von 1963, 1966 und 1967 zusammen. Diese seien daher nur in kurzen Worten charakterisiert. Falls notwendig, wird an entsprechender Stelle noch einmal näher auf sie eingegangen. BUCHWALD (1963) stellt erstmals den Grundgedanken, nämlich die Ersetzung des Begriffs "reinforcement" durch gedächtnispsychologische Begriffe und Entscheidungsregeln, in mehreren mathematisch formulierten Modellvarianten vor. BUCHWALD (1966) untersucht nochmals verschiedene Methoden zur Herstellung eines empirischen Wiederholungsniveaus nach b. Seine Daten unterstützen die Einführung des Parameters g (s.u.). In BUCHWALD (1967) wird geprüft, inwieweit sich die Effekte "verzögerter" Rückmeldung in das Modell einpassen.

BUCHWALD (1969) bezieht sich explizit auf die Experimente THORNDIKES (1932), die kurz durch $k > 2$ und fehlende Wirksamkeit von f gekennzeichnet sind. In Abhebung von dem asymmetrischen "revidierten Gesetz des Effekts" bietet er eine "neue Interpretation" an, die folgendes leisten soll: (a) Die Wirkung von r wird erklärt ohne Bezugnahme auf Begriffe wie Triebreduktion, Verstärkung usw.; (b) Die Asymmetrie zuungunsten von f kann erklärt werden, ohne unterschiedlich starke Wirkung von r bzw. f auf S-R-Verbindungen annehmen zu müssen; (c) Angabe der Bedingungen, unter denen die Wirkungen von r und f vergrößert oder verkleinert werden.

Der Grundgedanke der "neuen Interpretation" ist folgender: Wenn im Lerndurchgang die Sequenz Reizwort S - Reaktion R - Rückmeldung (outcome) O auftritt, bilden sich per Kontiguität getrennt und unabhängig voneinander S-R- bzw. S-O-Verbindungen, und diese beiden Assoziationen bestimmen im Prüfdurchgang die Leistung der Vp.

Die Begründung für dieses Postulat der unabhängigen S-R-bzw. S-O-Assoziationen entnimmt BUCHWALD u.a. den Berichten von Vpn aus Experimenten von DAND (1946) und BOWER (1962), in

denen die Vpn erklärten, sie hätten sich zwar im Prüfungsdurchgang noch an ihre eigene Antwort aus dem Lerndurchgang erinnern können, aber nicht mehr an die Rückmeldung, und umgekehrt.

Der Schlüssel zum Verständnis der Asymmetrie von r und f zugunsten von f liegt in folgendem (falls die Asymmetrie in bezug auf das Zufallsniveau $1/k$ beurteilt wird):

1. Die Daten zur Bestimmung der relativen Effizienz von r und f sind Wiederholungshäufigkeiten.
2. Je häufiger nach r (bzw. nach f) die Reaktionen wiederholt werden, um so wirksamer (bzw. um so weniger wirksam) wird r (bzw. f) beurteilt.
3. Jeder Mechanismus, der, unabhängig von r bzw. f , das Wiederholen von Reaktionen begünstigt, verringert somit die Chance für f , als wirksam beurteilt zu werden.

4. Als einen solchen Mechanismus führt BUCHWALD hypothetisch die Möglichkeit ein, daß Vpn im Prüfungsdurchgang, wenn sie auf ein Reizwort S hin sich zwar an ihre Reaktion R , nicht mehr aber an die Rückmeldung O erinnern können, überzufällig häufig ihre eigene Reaktion wiederholen ("response bias").

BUCHWALD gibt keinen unmittelbaren Beweis für diese Hypothese, er schreibt vielmehr vorsichtig: "... It seems reasonable to believe..." (in BUCHWALD, 1966, S. 83; die Hypothese erhält dort gewisse experimentelle Unterstützung).

Ähnliche Überlegungen und Stützung findet man aber auch bei anderen Autoren (z.B. DAND, 1946; BERNBACH, 1965; NUTTIN u. GREENWALD, 1968, S. 61 u. 152; MARX, 1971).

Bei der obigen Erklärung der Asymmetrie (bezüglich $1/k$) spielen also nicht unterschiedliche Informationswerte und Strategien beim Lernen (Einspeichern) eine Rolle wie bei BUCHWALD (1962). Die Asymmetrie resultiert vielmehr erst aus den Gedächtnis- und Entscheidungsprozessen beim Wiederfinden und Verarbeiten von Informationen im Prüfungsdurchgang. Nichtsdestoweniger argumentiert BUCHWALD sowohl 1962 als auch 1969 mit Begriffen der Informationsverarbeitung (in Entgegensetzung zum THORNDIKEschen Standpunkt).

Einschränkend zu Experimenten, in denen Wiederholungshäufigkeiten zur Beurteilung der relativen Effizienz benutzt werden, schreibt BUCHWALD (1969, S. 134) die Punkte 1. bis 4. zusammenfassend: "From this viewpoint the asymmetry resides in the general experimental situation rather than in the effects of 'Right' and 'Wrong' on the strength of associative connections" (Die Berücksichtigung dieser Einschränkung wird später zur Einführung eines relativen Effizienzmaßes führen, das nicht aus Wiederholungshäufigkeiten abgeleitet ist; s. 6.3.4.)

Das jetzt vorzustellende probabilistische Modell ist für eine modifizierte Form der THORNDIKE-Experimente entwickelt worden, wie sie schon im Zusammenhang mit der IT-Prozedur auf S. 21 beschrieben wurde. Die Modellaxiome sind (BUCHWALD, 1969, S. 134f.):

- "A1. On a test trial the probability of recalling the previous response to a given stimulus does not depend on the outcome that followed that response. Let c be the probability that a response will be recalled.
- A2. On a test trial the probability of recalling the outcome that followed the previous response to a given stimulus item does not depend on the outcome that occurred. Let u be the probability that an outcome will not be recalled, and $1-u$ the probability that an outcome will be recalled.
- A3. On a test trial the probability of repeating the response to a given stimulus depends upon whether or not the response is recalled, and upon whether or not its outcome is recalled.
 - a) When a previous response is not recalled the probability that it will be repeated equals the probability of its being made on the information trial.
 - b) When a previous response is recalled:
 - i) the probability that it will be repeated is one, given that the outcome was 'Right' and the outcome is recalled;
 - ii) the probability that it will be repeated is zero, given that the outcome was 'Wrong' and the outcome is recalled;
 - iii) the probability that it will be repeated has some value, g , if the outcome is not recalled, or if the outcome was nothing (even if the outcome is recalled)."

Um den Forderungen der Axiome möglichst zu entsprechen, wurden in der experimentellen Anordnung diesbezügliche Maßnahmen

getroffen. Für die angenäherte Gültigkeit von A1 und auch A2 ist es erforderlich, daß zwischen Rückmeldung O und nächstem Reizwort S keine Zeit für "stilles, inneres Wiederholen" (rehearsal) bleibt. Spezielle Strategien zugunsten von r bzw. f (s. BUCHWALD, 1962) sollten also ausgeschaltet sein. Im weiteren soll unter Rehearsal nicht nur das mechanische stille Wiederholen von z.B. S-R-Verbindungen verstanden werden, sondern im weiteren Sinne auch das zweckmäßige Kodieren mit Hilfe von "Eselsbrücken".

Entsprechend A2 sind die Rückmeldungen r, f und b gleich häufig und zufällig gemischt im Lerndurchgang vorgesehen, damit nicht durch etwaige Isolationseffekte die eine oder andere Rückmeldung bevorzugt eingeprägt und erinnert wird.

Um die Gleichwahrscheinlichkeit der Reaktionen nach A3a angenähert zu gewährleisten, werden als Antwortalternativen Ziffern (z.B. 1,3,5,7) und als Reizworte häufig auftretende bekannte Wörter benutzt, um so starke präexperimentelle Bevorzugungen und Assoziationen auszuschalten. Es werden ungerade Ziffern benutzt, um Reihenfolgeeffekte möglichst zu eliminieren.

Ein näherer Kommentar ist zu A3b(iii) erforderlich: Wenn also die Reaktion R aus dem Lerndurchgang von der Vp erinnert wird, die Vp sich aber nicht mehr an O erinnert oder sich daran erinnert, daß O=b war, dann wiederholt sie die Reaktion mit der Wahrscheinlichkeit g. Derselbe Parameter g dient also zu zwei Zwecken:

Einmal spiegelt er eine gewisse Tendenz wieder, falls O vergessen wurde, auf der einmal gegebenen Antwort zu beharren (oder, was auf das Gleiche hinausläuft, bevorzugt auf O=r zu tippen), das andere Mal beschreibt er, wie die Vpn überhaupt auf b reagieren. Das kann z.B. stark von der Einführung von b in der Instruktion abhängen. (In späteren Diskussionen wird g zeitweilig in g' bezüglich Vergessen von O und in g'' bezüglich b aufgespalten, um sich klar auf die beiden verschiedenen Zwecke beziehen zu können.)

BUCHWALD (1963, S. 5) sagt dazu: "This is a strong assumption which is made here more because it avoids adding a

fourth parameter than because it seems likely to be true." - Innerhalb der Versuchsanordnung und des Modells von BUCHWALD besteht jedenfalls keine Möglichkeit, diese beiden Aspekte von g zu trennen (s. dagegen die Ergebnisse im Hauptexperiment unter 7.1.3).

Der Parameter g spielt eine entscheidende Rolle in der obigen Argumentation (S. 28, Punkte 1. bis 4.) bezüglich der relativen Effizienz von r und f zuungunsten von f , falls sich auf das Zufallsniveau $1/k$ bezogen wird. Der unter Punkt 3. angesprochene Mechanismus wird im weiteren als "g-Mechanismus" bezeichnet werden. Seine genauere Wirkungsweise soll erst veranschaulicht werden, nachdem aus den Axiomen A1 bis A3 leichter überschaubare Gleichungen für die theoretischen Wiederholungswahrscheinlichkeiten nach r , f und b abgeleitet worden sind.

Die Axiome A1 bis A3 lassen sich getrennt nach r , f und b graphisch in Form von "Baumdiagrammen" darstellen (s. Abb. 2.3.3-1).

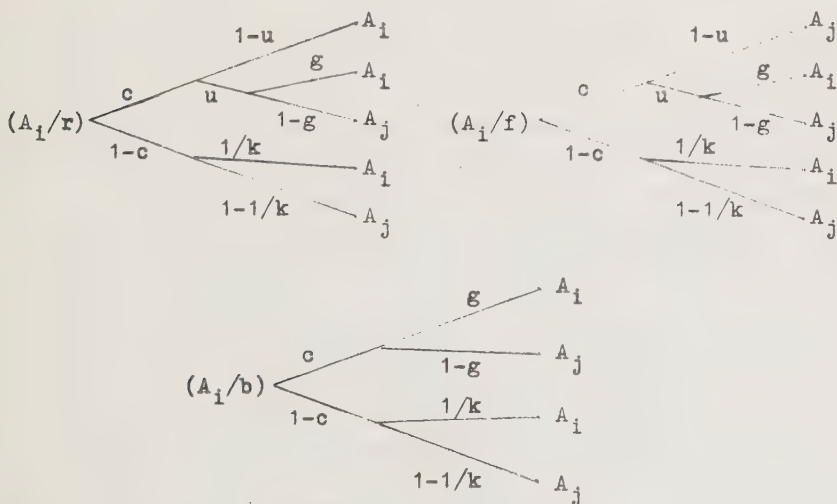


Abb. 2.3.3-1: Zur Erläuterung s. Text

Die oben dargestellten Diagramme werden im folgenden BUCHWALD- bzw. kürzer B-Diagramme genannt werden. Zur Erläuterung der Abb. 2.3.3-1: Am äußersten linken Knoten stehen die Reaktionen A_i , die im Lerndurchgang gegeben wurden, und die zugehörigen Rückmeldungen r , f oder b . Nach rechts folgen dann die Verzweigungen bezüglich Erinnerung oder Nichterinnerung an Reaktionen ($c; 1-c$) und Rückmeldungen ($1-u; u$) und die Verzweigungen bezüglich des Antwortverhaltens bei unzureichenden Erinnerungen oder bei Erinnerung an b ($1/k; 1-1/k; g; 1-g$). Die Symbole A_i rechts am Ende der einzelnen "Erinnerungs- und Entscheidungswege" besagen, daß die Antworten A_i aus dem Lerndurchgang im Prüfdurchgang wiederholt wurden, die Symbole A_j besagen, daß die Antworten A_i nicht wiederholt wurden.

Aus den B-Diagrammen (d.h. also aus den Axiomen) lassen sich leicht die theoretischen Wiederholungswahrscheinlichkeiten nach r , f und b ableiten. Dabei bedeutet $P(A_i/O)$ die Wahrscheinlichkeit, daß A_i im Prüfdurchgang wiederholt wird, falls auf A_i im Lerndurchgang die Rückmeldung O ($= r$ oder f oder b) gefolgt war:

$$P(A_i/r) = c \cdot (1-u) + c \cdot u \cdot g + (1-c)/k \quad (1a)$$

$$P(A_i/f) = c \cdot u \cdot g + (1-c)/k \quad (1b) \quad ("R-corr")$$

$$P(A_i/b) = c \cdot g + (1-c)/k \quad (1c)$$

(Im später zu referierenden Experiment I wird auf diese Gleichungen unter der Abkürzung "R-corr" Bezug genommen.)

Ausführlich liest sich z.B. die Gleichung für $P(A_i/f)$ so: Falls im Lerndurchgang auf A_i die Rückmeldung $O=f$ folgte, wird A_i im Prüfdurchgang wiederholt, weil entweder die V_p sich an A_i mit (der Wahrscheinlichkeit) c erinnert und f mit u vergessen hat und nun mit g auf A_i beharrt oder die V_p A_i mit $(1-c)$ vergessen hat und nun mit $1/k$ wieder A_i sagt.

Anhand der Gleichung für $P(A_i/f)$ kann nun auch die Wirkung des g -Mechanismus demonstriert werden. BUCHWALD hat die Struktur seiner Axiome bzw. der aus ihnen abgeleiteten Gleichungen (1a) bis (1c) so gewählt, daß bei bestimmten Werten der Parameter die empirisch oft gefundene Ineffizienz von f erklärt werden kann. Diese Ineffizienz drückt sich z.B.

darin aus, daß die Wiederholungshäufigkeit nach f sogar über Zufallswiederholungsniveau $1/k$ liegen kann, sich also noch nicht einmal ein geringfügiger Unterdrückungseffekt zeigt (s. z.B. auf S. 13 die Ergebnisse THORNDIKES: $29\% > 20\%$). Falls also die Gleichung (1b) überzufälliges Wiederholen nach f beschreiben soll, heißt das:

$$P(A_i/f) = c \cdot u \cdot g + (1-c)/k > 1/k.$$

Diese Ungleichung kann vereinfacht werden:

$$c \cdot u \cdot g - c/k > 0; \text{ und weiter: } u \cdot g > 1/k.$$

(Daraus folgt auch noch: $g > 1/k$, da ja $u < 1$.)

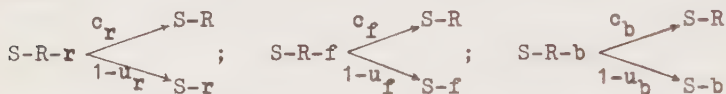
Wenn also die Parameter u und g so große Werte haben, daß ihr Produkt $> 1/k$ ist, dann kann überzufälliges Wiederholen nach f , d.h. eine starke Ineffizienz von f , erklärt werden. Es muß also (1.) u = Wahrscheinlichkeit für Vergessen von O groß sein, denn nur dann kann ja der g -Mechanismus überhaupt und häufig genug in Aktion treten, (2.) muß auch g groß genug sein, also die Tendenz, auf der erinnerten Antwort A_i zu beharren.

Man sieht, daß durch den additiven Term $c \cdot u \cdot g$ in (1b) die Effizienz von f gesteuert wird. Falls u und g kleinere Werte annehmen, muß sich nicht unbedingt mehr $P(A_i/f) > 1/k$ ergeben; auf jeden Fall aber trägt der Term $c \cdot u \cdot g$ zur Erhöhung des Wiederholungsniveaus, also zur Verringerung der Effizienz von f bezüglich $1/k$ bei.

Da eine so erklärte Asymmetrie zuungunsten von f bezüglich $1/k$ von einem an sich problemirrelevanten Mechanismus, verbunden mit der speziellen experimentellen Situation (s. S. 29), abhängt, war es nützlich, mit Hilfe der b -Rückmeldung diesen Effekt zu neutralisieren, es jedenfalls zu versuchen. Das Wiederholungsniveau nach b enthält ja auch den g -Mechanismus. (Dieses kann man sich anhand des leicht modifizierten B-Diagramms (A_i/b) auf S. 68 klarmachen. Dort ist der Parameter g in g' und g'' aufgespalten: g' bezieht sich auf die Erinnerung an b , g'' auf das Vergessen von O ; man müßte also genauer vom g'' -Mechanismus sprechen.) Bezieht man jetzt die Beurteilung der Effizienz von r und f nicht mehr auf $1/k$,

sondern auf das Wiederholungsniveau nach b, so fällt der Einfluß des g- (genauer g"-)Mechanismus weg.

In den Axiomen bzw. in den daraus abgeleiteten Gleichungen (1a) bis (1c) sind drei Unabhängigkeitspostulate enthalten, die der Deutlichkeit halber herausgestellt werden sollen. Die folgenden schematischen Zeichnungen dazu beziehen sich auf die Erinnerung an die S-R- und S-O-Verbindungen im Prüfdurchgang:



Die Unabhängigkeitsforderungen nach A1 und A2 lauten kurz:

$$U1: c_r = c_f = c_b = c \quad (\text{nach A1})$$

$$U2: 1-u_r = 1-u_f = 1-u_b = 1-u \quad (\text{nach A2})$$

$$U3: P(\langle \text{S-R} \rangle \cap \langle \text{S-O} \rangle) = c \cdot (1-u) \quad (\text{Forderung der separaten, unabhängigen Gedächtnisprozesse})$$

Die Schreibweise in U3 besagt: Die Wahrscheinlichkeit, bei Darbietung von S sich sowohl an R als auch an O zu erinnern, ist gleich dem Produkt der Einzelerinnerungswahrscheinlichkeiten.

BUCHWALD führt nun einige speziell geplante Tests durch, um die postulierten Grundannahmen einzeln zu überprüfen. Auf genauere Darstellung der Experimente soll hier verzichtet werden, da sich die Voruntersuchungen (s. Kapitel 5) sehr stark an diese Versuchsanordnungen anlehnen und dort näher beschrieben werden. - Im folgenden tauchen spezielle Bezeichnungen für die verschiedenen Versuchsbedingungen auf. Da es sich z.T. nicht um allgemeine Fachausdrücke handelt, sollen sie der Einfachheit halber unübersetzt bleiben.

(1) "reminder condition": Der Gedächtnisprozeß hinsichtlich der Erinnerung an R (d.h. an A_1) soll eliminiert werden. Dies geschieht dadurch, daß die Vp im Prüfdurchgang an ihre früher im Lerndurchgang gegebene Antwort explizit erinnert ("reminder") wird. Sie braucht sich also dann nur noch an O zu

erinnern. Der Parameter c erhält also aufgrund der experimentellen Manipulation den Wert $c = 1$. Die Gleichungen (1a) bis (1c) gehen über in:

$$P_{rc}(A_i/r) = (1-u) + u \cdot g \quad (2a)$$

$$P_{rc}(A_i/f) = u \cdot g \quad (2b) \quad \begin{array}{l} \text{"R-remind";} \\ \text{s.Experiment I} \end{array}$$

$$P_{rc}(A_i/b) = g \quad (2c)$$

Der Index rc bei P_{rc} bedeutet "reminder condition".

Aus den Gleichungen ergibt sich eine Vorhersage:

Bildet man die Differenzen

$$(1a)-(1b) = c \cdot (1-u) \quad \text{("nonreminder"-Bedingung) und}$$

$$(2a)-(2b) = 1-u \quad \text{("reminder"-Bedingung),}$$

so sieht man, daß die "reminder"-Bedingung den Effekt der Rückmeldungen r und f vergrößern sollte ($c \cdot (1-u) < (1-u)!$).

Verstärkungstheorien dagegen machen die Voraussage, daß unter der "reminder"-Bedingung der Unterschied in den Wiederholungshäufigkeiten nach r bzw. f verschwinden sollte.

Im Experiment waren in derselben Liste Wörter der "reminder"- und der "nonreminder"-Bedingung gleich häufig. Die V_{pn} wurden zwischen dem Lern- und dem Prüfdurchgang instruiert, daß ihnen im Prüfdurchgang bei einigen Reizworten gesagt würde, welche Antworten sie zu eben diesen Reizworten im Lerndurchgang genannt hätten. Die Ergebnisse entsprachen der Tendenz nach der Vorhersage des BM.

Die "reminder"-Bedingung bietet einen weiteren Test an: Aus den Daten erhält man nach (2c) und (2b) Schätzungen $\hat{u}$ und $\hat{g}$ von u und g . Setzt man $\hat{u}$ und $\hat{g}$ in (2a) ein, ergibt sich eine quantitative Voraussage für $P_{rc}(A_i/r)$. Aus den Wiederholungshäufigkeiten der r -Items hat man aber auch schon eine andere Schätzung, $\hat{P}_{rc}(A_i/r)$. Es müßten $(1-\hat{u}) + \hat{u} \cdot \hat{g}$ und $\hat{P}_{rc}(A_i/r)$ bis auf Zufallsschwankungen übereinstimmen. BUCHWALD findet befriedigende Übereinstimmung.

Zur Beurteilung des Werts dieses Tests sind zwei Anmerkungen zu machen: a) Die Gleichungen (2a) bis (2c) enthalten von den Unabhängigkeitsforderungen $U1$ bis $U3$ nur $(1-u_r) = (1-u_f)$ aus

U2. b) Die Voraussetzungen für die Anwendung des χ^2 -Tests beim Vergleich von $(1-\hat{u})+\hat{u} \cdot \hat{g}$ und $\hat{P}_{rc}(A_i/r)$ sind nicht ganz sauber erfüllt.

(2) "delayed outcome": Der Gedächtnisprozeß hinsichtlich der Erinnerung an O soll eliminiert werden. Dies geschieht, nicht ganz analog zur "reminder"-Bedingung, dadurch, daß die Vp im Lerndurchgang keine Rückmeldung erhält, im Prüfdurchgang aber nach Darbietung von S erfährt, ob ihre früher gegebene Antwort richtig oder falsch war. Sie braucht sich also nur noch an R zu erinnern. Der Parameter 1-u erhält also aufgrund der experimentellen Manipulation den Wert 1-u=1 (bzw. u=0). Die Gleichungen (1a) bis (1c) gehen über in:

$$P_{do}(A_i/r) = c + (1-c)/k \quad (3a)$$

$$P_{do}(A_i/f) = (1-c)/k \quad (3b) \quad \begin{matrix} \text{"O-remind";} \\ \text{s.Experiment I} \end{matrix}$$

$$P_{do}(A_i/b) = c \cdot g + (1-c)/k \quad (3c)$$

Der Index do bei P_{do} bedeutet "delayed outcome". - Die Gleichung (3c) wird von BUCHWALD selbst nicht angeführt; vermutlich, weil wegen u=0 der g-Mechanismus nicht betrachtet zu werden braucht. Hier bestünde jedoch die Möglichkeit eines quantitativen Tests wegen der Gleichheit von (1c) und (3c). Auch die simultane Schätzmöglichkeit bezüglich c aus (3a) und (3b) wird nicht diskutiert.

Aus den Gleichungen ergibt sich eine Vorhersage:

Bildet man die Differenzen

$$(1a)-(1b) = c \cdot (1-u) \quad \text{"immediate outcome"-Bedingung) und}$$

$$(3a)-(3b) = c \quad \text{"delayed outcome"-Bedingung),}$$

so sieht man, daß die "delayed outcome"-Bedingung den Effekt von r bzw. f vergrößern sollte ($c > c \cdot (1-u)$!). Verstärkungstheorien dagegen sagen bei verzögerter Verstärkung eher schlechtere Leistung voraus.

In einem Experiment (BUCHWALD, 1967) waren in der Lernliste "immediate outcome"-, "delayed outcome"- und "no information"-Items gemischt. Bei den "delayed outcome"-Items sagte der V1 im Lerndurchgang nichts; er informierte bei diesen Items die

Vpn erst im Prüfdurchgang, ob ihre Reaktionen im Lerndurchgang richtig oder falsch gewesen waren. Aufgrund einiger Komplikationen führte BUCHWALD (1969) noch eine modifizierte Version der "delayed outcome"-Bedingung durch; auf Einzelheiten kann verzichtet werden. Die gemeinsamen Ergebnisse aus den Arbeiten (1967, 1969) sprechen für das BM und gegen die Verstärkungshypothese.

Die beiden Testkomplexe "reminder condition" und "delayed outcome" beziehen sich nur indirekt auf die Unabhängigkeitsannahme U3 (Forderung separater, unabhängiger Gedächtnisprozesse; s. S. 34). Ein direkter Test von U3 kann erst im Anschluß an das Experiment III (s. 7.1) mit Hilfe der dort angewandten Methode ("Verbalisieren") prinzipiell durchgeführt werden.

(3) "apparent recall": Entsprechend dem BM werden nicht alle Antworten, die im Lerndurchgang gegeben worden waren und im Prüfdurchgang erinnert werden, wiederholt (falls z.B. gleichzeitig sich an S-f erinnert wird). Führt man nun einen "recall"-Prüfdurchgang ein, in dem die Vpn aufgefordert werden, möglichst alle ihre Antworten aus dem Lerndurchgang zu sagen, unabhängig davon, ob sie von r, f oder b gefolgt waren, so ist folgendes zu erwarten: Die Wahrscheinlichkeit für diese "apparent recalls" wird höher sein als für die bisher benutzten, evtl. durch r, f und b beeinflussten Antwortwiederholungen. Dieser Effekt sollte unter r am geringsten, unter f am größten sein.

Für die Wahrscheinlichkeit ergibt sich aus dem BM, in dem die Antwortentscheidungsregel entsprechend abgeändert ist:

$$P(\text{"apparent recall"}) = c + (1-c)/k \quad (4) \quad (\text{"R-recall"; s.Experiment I})$$

Wie man leicht sieht, ist diese Erinnerungswahrscheinlichkeit größer als alle Wiederholungswahrscheinlichkeiten nach (1a) bis (1c). Nach der Unabhängigkeitsannahme U1: $c_r = c_f = c_b$ gilt die Gleichung (4) für die r-, f- und b-Items getrennt, aber auch für alle zusammen.

Im Experiment wurden sowohl die "konventionellen" Prüfdurchgänge als auch "recall"-Durchgänge nacheinander durchgeführt. Die Resultate unterstützten die obige Voraussage. Besonders wichtig ist das Teilergebnis aus den "recall"-Daten, daß die Reaktionen R nach r, f und b gleich gut behalten wurden. Das ist eine Stützung der Unabhängigkeitsforderung U1: $c_r = c_f = c_b$! BUCHWALD (1969; Tabelle 4) hat die diesbezüglichen Ergebnisse über $k=2,3,4$ und 6 gemittelt dargestellt. Es ist so nicht zu ersehen, ob U1 für verschiedene k unterschiedlich gut erfüllt ist.

Im "recall"-Durchgang wurden die Vpn auch nach den Rückmeldungen O aus dem Lerndurchgang gefragt ("O-recall"; s. Experiment I). Leider gibt BUCHWALD dazu keine Ergebnisse an. Hier hätte die Möglichkeit bestanden, auch die Unabhängigkeitsforderung U2: $(1-u_r) = (1-u_f) = (1-u_b)$ nachzuprüfen.

BUCHWALD diskutiert noch einige andere Arbeiten (z.B. POSTMAN u. ADAMS, 1955; GREENWALD, 1966), in denen auch "recall"-Daten erhoben wurden: Daß sich dort eine signifikante Überlegenheit der "apparent recalls" nach r über die nach f ergibt, wird mit unterschiedlichem Rehearsal erklärt. Der Effekt unterschiedlichen Rehearsals kann aber bei BUCHWALDs Experiment nicht (stark) auftreten, da ja hier die Gelegenheit bzw. Zeit zum "stillen, inneren Wiederholen" minimiert war.

Da das Experiment (3) als Daten sowohl Wiederholungshäufigkeiten ("repetition measures") als auch Erinnerungsdaten ("recall measures") liefert, ist es nach BLÖSCHL (1969, S. 39) unter Wiederholungs- und unter Behaltensprüfungen zu klassifizieren. Dieser Hinweis ist nicht trivial; s. dazu unter 6.3.4.

Abschließend sagt BUCHWALD zu den Tests (1) bis (3) bezüglich seines Modells: "In general the results were in accord with theoretical predictions." Er betont aber auch den eingeschränkten Rahmen seines Ansatzes. Dieser ist ja hauptsächlich auf eine Reinterpretation der THORNDIKE-Experimente zugeschnitten, die ihrerseits auch nur beschränkte Aussagen ermöglichen (s. POSTMAN, 1962, S. 360).

Zum Schluß soll noch das Problem der Parameterschätzungen für c , u und g behandelt werden. Es ist auffallend, daß BUCHWALD (1969) keine direkten Angaben dazu macht, obwohl mindestens unter den Spezialbedingungen(1) bis (3) Parameterschätzungen leicht möglich wären.

In BUCHWALD (1963) werden c , u und g mit Hilfe einer Minimum- χ^2 -Prozedur geschätzt, allerdings für eine etwas abgeänderte experimentelle Anordnung (PIT). Die Parameter u und g sind danach im wesentlichen unabhängig von k , der Parameter c hängt dagegen stark von k ab; BUCHWALD (1969) geht bei Darstellung seines Modells nicht mehr darauf ein: k scheint in den Gleichungen nur durch den Zufallsanteil $(1-c)/k$ ins Spiel zu kommen.

Einschränkend ist zu sagen, daß die Voraussetzungen für die Durchführung der Minimum- χ^2 -Prozedur und die nachfolgenden Signifikanztests in der Arbeit von 1963 nicht sauber erfüllt waren (s. dazu ATKINSON et al., 1965, S. 394). Auch auf die Verschiebung des nominalen Signifikanzniveaus bei Durchführung einer großen Zahl von Tests - bei BUCHWALD (1963) sind es 320 χ^2 -Tests - sei hingewiesen (s. z.B. SACHS, 1969, S. 496).

Trotz einiger experimenteller und theoretischer Unzulänglichkeiten stellt das BM einen vielversprechenden Ansatz dar. Es wird in den Voruntersuchungen (s. Kapitel 5) versucht werden, die Experimente und Ergebnisse BUCHWALDs (1969) zu reproduzieren, eventuell auch auszuweiten. Dabei sollen einige der erwähnten Unzulänglichkeiten möglichst eliminiert werden.

2.4 Arbeiten im Anschluß an BUCHWALD (1969)

In diesem Abschnitt sollen einige Anmerkungen gemacht und solche Arbeiten zitiert werden, die in mehr oder weniger starker Beziehung zu dem BM (1969) stehen. Sie sind aus 2.3.3 herausgenommen worden, um die dortige Darstellung des BM nicht zu sehr zu belasten.

Bemerkenswert ist, daß es in der Literatur keine Untersuchung gibt, die das BM als mathematisches Modell in der ursprünglichen oder auch in abgeänderter Form noch einmal behandelt (zur einzigen Ausnahme s.u. HUMPHREYS, 1972). Der Grundgedanke BUCHWALDS, also das Postulat zweier Gedächtnisprozesse bezüglich S-R und S-O, jedoch wird häufiger zitiert (z.B. GREENWALD, 1970; ATKINSON u. WICKENS, 1971; ESTES, 1971; WITTER et al., 1971; LIU, 1972).

Die besondere Schwierigkeit beim Vergleich mit anderen Untersuchungen liegt darin, daß in diesen im allgemeinen Rehearsal gestattet ist. Die Verhinderung von Rehearsal ist aber gerade von zentraler Bedeutung für die Gültigkeit der Unabhängigkeitsforderungen U1 bis U3.

Eine Ausnahme stellt die Arbeit von LIU (1972) dar, in der versucht wurde, die Zeit zwischen Rückmeldung und nächstem Reizwort zu minimieren. (Es wurde nur $k=2$ behandelt.) Neben einigen anderen Abänderungen der THORNDIKEschen bzw. BUCHWALDSchen Versuchsanordnung ist besonders folgender Aspekt wichtig: In einer Gruppe (RW) wurde, wie üblich, mit "Right" und "Wrong" rückgemeldet. In einer anderen Gruppe (AN) wurde nach jeder Antwort den Vpn vom V1 "Army" oder "Navy" gesagt. Die Vpn erfuhren erst nach dem Lerndurchgang, aber vor dem Prüfdurchgang, daß "Army" z.B. "Right" bedeutete und "Navy" "Wrong". In einer weiteren Versuchsbedingung wurden die Vpn schon vor dem Lerndurchgang über die Bedeutung von "Army" bzw. "Navy" aufgeklärt (AN-RW-Gruppe).

Nach dem BM dürften sich keine Unterschiede zwischen diesen Gruppen ergeben, da a priori Kontiguitätsassoziationen zwischen Reizwörtern S und Rückmeldungen wie "Right", "Wrong", "Army" und "Navy" sich in gleicher Weise bilden und erinnern lassen sollten.

Die Ergebnisse zeigen, daß im wesentlichen nur die Rückmeldung "Right" in der RW-Gruppe einen signifikanten Effekt erbrachte, sowohl was die Wiederholungshäufigkeiten als auch die Erinnerung an die Rückmeldungen anbetrifft. LIU schließt daraus, daß im BM die Rolle des differentiellen Rehearsals

noch unterschätzt wird. Offensichtlich war in seinem Versuch die Verhinderung von Rehearsal so wirksam, daß den Vpn gerade noch die Zeit blieb, aufgrund einer eingeschliffenen Gewohnheit ("established habit") Sequenzen von der Form S_i-R_j -Right (bzw. nur S_i-R_j) effektiv im Stillen zu wiederholen. Es blieb aber nicht Zeit, nach "Wrong" eventuell auf die andere Antwort umzuschalten oder sich über die jeweilige Bedeutung von "Army" bzw. "Navy" klarzuwerden.

Die Folgerung aus dieser Untersuchung ist in gewisser Übereinstimmung mit der später in vorliegender Arbeit vertretenen Meinung bezüglich automatisierter Strategien (s. 6.3.3). Als einzige Arbeit zu diesem Thema wurde sie aber zunächst noch nicht zu stark gewichtet: Es sollten erst die Ergebnisse eigener Experimente abgewartet werden. - Im Prinzip ist die Versuchsanordnung LIUS übrigens geeignet, die Genese der Wirkung von Belohnung und Bestrafung im menschlichen Lernen zu untersuchen.

Eine Arbeit, die dem Ansatz BUCHWALDS (1969) sehr stark ähnelt, ist die von HUMPHREYS (1972). Seine experimentelle Anordnung und das mathematische Modell dazu sind auf etwaige "bias"-Prozesse ausgerichtet, die möglicherweise eine Asymmetrie zuungunsten von f bewirken, falls die Rückmeldung im Prüfdurchgang nicht mehr erinnert werden kann. Im BM ist der Ort dieser Asymmetrie bei dem "bias"-Parameter g zu suchen. (Im Augenblick wird nur der g -Aspekt von g diskutiert, der sich auf das Vergessen von O bezieht; s. S. 30.)

Die Bedeutung von g (genauer g') war erklärt worden entweder als Beharrung auf der einmal gegebenen Antwort oder als Rate-"bias" zugunsten von r . Diese beiden möglichen Bedeutungen von g' , die im BM nicht unterschieden sind, werden bei HUMPHREYS sowohl experimentell als auch theoretisch separiert, und es wird gezeigt, daß daraus ein Vorteil zugunsten der Effizienz von r resultiert.

Die Asymmetrie wird durch den Unterschied in den Fehlerhäufigkeiten bestimmt; es werden also keine Wiederholungshäufigkeiten bezüglich irgendeiner neutralen Bezugslinie (z.B. bb)

beurteilt.

Ähnlich wie BUCHWALD versucht HUMPHREYS, mit Hilfe von Entscheidungsregeln (entsprechend A3 im BM) die Erinnerungen an die Ereignisse im Lerndurchgang (S-R- und S-O-Assoziationen im BM) mit den Antworten im Prüfdurchgang zu verbinden. Die Anpassung seines Modells an die Daten ist allerdings nicht allzu gut. Nach den üblichen Kriterien eines statistischen Tests wäre das Modell sogar auf dem 5%-Niveau abzulehnen (s. aber zur Beurteilung von mathematischen Lernmodellen BUSH, 1963). Von fünf Parametern sind nur drei identifizierbar (zum Problem der Identifizierbarkeit von Parametern s. GREENO u. STEINER, 1964; RESTLE u. GREENO, 1970).

Zu diesen Modellproblemen kommt auf experimenteller Seite hinzu, daß wegen der Modifikation des "normalen" Richtig-Falsch-Paradigmas die Vergleichbarkeit mit anderen Arbeiten zu diesem Gebiet, einschließlich BUCHWALDs, sehr eingeschränkt und die Ableitung von Hypothesen zu Latenzzeiten erschwert ist. Trotz des interessanten Ansatzes soll daher das Modell und die Versuchsanordnung HUMPHREYS' nicht weiter betrachtet werden.

Mangels weiterer Literatur, die sich unmittelbar auf das BM bezog, war es naheliegend zu versuchen, das BM in die schon zahlreich vorhandenen mathematischen Lernmodelle zu integrieren (s. dazu die Einführung in die Mathematische Lerntheorie von ATKINSON et al., 1965). Zwei Gruppen von Modellen boten sich an: Die eine behandelt explizit das Eliminieren von falschen Antworten unter der Nichtkorrektur-Methode, die andere befaßt sich mit Vermeidungslernen, speziell also mit dem Lernen nach Schock (Flucht: Mißerfolg) bzw. Nicht-Schock (Vermeidung: Erfolg).

Zur ersten Gruppe gehören u.a. die Arbeiten von BOWER (1962), NAHINSKY (1964, 1966, 1967), NAHINSKY u. MUELLER (1968). BOWER (1962) macht bezüglich der Wirksamkeit von r und f von vornherein die Annahme, daß f keine Wirkung hat, wobei er auf die Ergebnisse von THORNDIKE (1932) hinweist. Da er sein

Modell nur in $k > 2$ -Situationen getestet, kommt er auch zu befriedigenden Ergebnissen. NAHINSKY (1964, 1966) führt für das Einprägen nach r und für das Eliminieren nach f zwei Parameter, c und a , ein, setzt sie aber zunächst gleich, um die mathematischen Probleme, besonders die der Parameterschätzung, lösen zu können; es wird also symmetrische Wirkung von r und f vorausgesetzt (s. dazu auch MILLWARD, 1964b). NAHINSKY (1967) und NAHINSKY u. MUELLER (1968) erlauben schließlich unterschiedliche Parameterwerte für c und a , die es gestatten, Aussagen über die relative Effizienz von r und f zu machen. Die Modelle sind aber noch nicht befriedigend und vor allem mathematisch so komplex, daß vorläufig verzichtet wurde, noch zusätzlich Vergessensprozesse einzuführen, wie sie ja im BM vorgesehen sind.

Zur zweiten Gruppe (Vermeidungslernen) gehören u.a. BUSH u. MOSTELLER (1955, S. 237f.), BOWER u. THEIOS (1964), THEIOS (1971). Beim üblichen Vermeidungslernen handelt es sich wohl-gemerkt um eine "correction"-Prozedur, da schließlich jeder Versuch mit der richtigen Reaktion endet, entweder durch Flucht oder Vermeidung. Neben anderen Parametern sind besonders die zwei wichtig (z.B. in THEIOS, 1971, e und s), die den Effekt eines Erfolgs- (Vermeidungs-) bzw. eines Fehler- (Flucht-) Durchgangs beschreiben sollen. Es zeigt sich nun aber, daß die Schätzung der Parameter, die also etwas über die relative Effizienz von Erfolgen bzw. Mißerfolgen aussagen könnten, noch sehr problematisch ist. Je nach der Schätz-methode bzw. nach dem Modell, von dem man ausgeht, scheint man zu gegensätzlichen Aussagen zu kommen (s. dazu ATKINSON et al., 1965, S. 299; HOLMAN, 1969; COOMBS et al., 1970, S. 278).

Die Identifizierbarkeit, genauer die Nichtidentifizierbarkeit der Parameter bereitet ebenfalls Schwierigkeiten: Nach THEIOS (1971, S. 313) haben in diesem letzteren Falle die Parameter wenig psychologische Bedeutung. Auch diese zweite Gruppe hat also noch zu sehr mit ihren eigenen Schwierigkeiten zu tun, als daß sie bei der Behandlung des BM hilfreich sein könnte.

Falls es keine Schätz- bzw. Identifizierbarkeitsprobleme gäbe, könnte man, wie schon gesagt, prinzipiell die beiden Parameterpaare c und a bzw. s und e aus den Modellen von NAHINSKY u. MUELLER (1968) bzw. THEIOS (1971) zur Bestimmung der relativen Effizienz positiver und negativer Verstärkung benutzen. Im BM genügt dazu der Parameter g , falls man ein relatives Effizienzmaß, bezogen auf das Wiederholungsniveau nach b , folgendermaßen definiert:

$$\text{relative Effizienz bezüglich } b = re(b) = \frac{P(A_1/r) - P(A_1/b)}{P(A_1/r) - P(A_1/f)}$$

Dabei sind die Wahrscheinlichkeiten $P(A_1/r)$ usw. aus den Gleichungen (1a) bis (1c) auf S. 32 zu entnehmen, und es ergibt sich

$$re(b) = \frac{c \cdot (1-u) + c \cdot u \cdot g + (1-c)/k - \langle c \cdot g + (1-c)/k \rangle}{c \cdot (1-u) + c \cdot u \cdot g + (1-c)/k - \langle c \cdot u \cdot g + (1-c)/k \rangle} = 1-g.$$

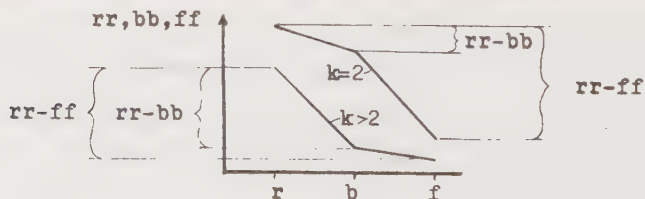
Genauer wäre zu schreiben: $re(b) = 1-g'$, wie man sich überzeugen kann, wenn man die kleine Rechnung anhand der modifizierten B-Diagramme auf S. 68 verfolgt.

Bei der praktischen Bestimmung von $re(b)$ sind natürlich $P(A_1/r)$, $P(A_1/f)$ und $P(A_1/b)$ durch rr , ff und bb zu schätzen, also:

$$re(b) = (rr-bb)/(rr-ff).$$

Es soll kein Unterschied in der Schreibweise zwischen dem theoretischen und dem geschätzten $re(b)$ gemacht werden.

BUCHWALD selbst hat kein solches Maß angegeben. Das vorgeschlagene Maß entspricht aber seiner Vorgehensweise, die Wirksamkeit von r bzw. f an der Differenz zum bb -Wiederholungsniveau zu beurteilen. Die Bedeutung von $re(b)$ sei an "typischen" Profilen zu $k=2$ und $k>2$ veranschaulicht:



Die Differenz $rr-ff$ ist ein allgemeines Maß für den durch r und f bewirkten Lernerfolg. Nimmt man nun bb als neutrales Bezugsniveau, dann ist $rr-bb$ der Anteil am Lernerfolg, der der Wirkung von r zuzuschreiben ist. $(rr-bb)/(rr-ff)$ ist dann ein Maß für die relative Wirksamkeit von r . Die relative Effizienz von f ist natürlich einfach das Komplement von $re(b)$ zu 100%.

Ohne weitere Rechnung ersieht man aus der obigen Skizze, daß für $k=2$: $re(b) < 50\%$ und für $k > 2$: $re(b) > 50\%$ ist. Es bedeutet also: $re(b) = 50\%$ Symmetrie von r und f ; $re(b) < 50\%$ Asymmetrie zugunsten von f ; $re(b) > 50\%$ Asymmetrie zugunsten von r . Dieses so erklärte Maß ist konsistent mit den Folgerungen BUCHWALDs (1962) aus seinen Daten. Es wird später mit einem anderen Maß für die relative Effizienz von r und f verglichen werden, das gewisse Nachteile von $re(b)$ vermeidet (s. 6.2.3 und 6.3.4).

2.5 Latenzzeiten beim BUCHWALD-Modell

Bevor in Kapitel 3 (Latenzzeiten: Stand der Forschung) näher auf empirische Befunde und Modelle zur Latenzzeitmessung eingegangen wird, soll kurz erläutert werden, welche Zeitdaten im Zusammenhang mit dem BM von Interesse sein werden. Diese Erläuterung wird Anhaltspunkte geben, warum im Kapitel 3 gerade die angegebenen Zeitdaten Aspekte besonders betont werden.

STERNBERG (1969a) gibt einen kurzen Überblick über die Annahmen, die bei der Analyse von (Reaktions-)Zeitverteilungen üblicherweise gemacht werden, wobei das Niveau der Komplexität von Mittelwerten über die Betrachtung von Varianzen und höheren Momenten bis zur näher spezifizierten Form der Zeitkomponentenverteilungen reicht. Da, wie noch zu zeigen ist, keine einschlägige Literatur existiert, die dazu berechnete, komplexere Hypothesen über Latenzzeiten bei der vorliegenden Untersuchung abzuleiten, wird die Analyse sich auf

dem Niveau von Mittelwerten bewegen; und zwar werden Mediane als Maß für zentrale Tendenzen benutzt werden (z.B. THOMAS, 1971).

Auf der Ebene der Wiederholungsdaten ist jede Vp oder jede Gruppe durch die Angabe von drei Werten gekennzeichnet, nämlich durch die relativen Wiederholungshäufigkeiten rr , ff und bb . Diese Größen entsprechen den "Wegen", die in den B-Diagrammen (S. 31) von der Antwort A_i im Lerndurchgang wieder zu A_i im Prüfdurchgang führen. Die Angabe der Nichtwiederholungshäufigkeiten $r\bar{r}$, $f\bar{f}$ und $b\bar{b}$ ist überflüssig; es ist ja $r\bar{r}=1-rr$, ..., $b\bar{b}=1-bb$.

Anders ist es bei der Analyse der Zeitdaten. Es ist durchaus möglich, daß die Zeiten, die über die "Wege" $A_i \rightarrow A_j$ zustandekommen, also zu den Größen $r\bar{r}$, $f\bar{f}$ und $b\bar{b}$ gehören, zusätzlichen Aufschluß über die zugrundeliegenden Gedächtnis- und Entscheidungsprozesse liefern. Jede Vp bzw. jede Gruppe wird also durch sechs Mediane gekennzeichnet werden, die folgendermaßen bezeichnet werden sollen: T_{rr} , $T_{r\bar{r}}$, T_{ff} , $T_{f\bar{f}}$, T_{bb} und $T_{b\bar{b}}$.

T_{rr} und $T_{f\bar{f}}$ sind "Erfolgs"-Latenzzeiten, $T_{r\bar{r}}$ und T_{ff} sind "Fehler"-Latenzzeiten. In der Literatur (s. Kapitel 3) wird oft von "success latencies" und "error latencies" gesprochen (z.B. HALL, 1969). Da es sich bei den einschlägigen Arbeiten fast ausnahmslos um PAL-Experimente unter der Korrekturbedingung handelt, sind die dort angegebenen "success"- bzw. "error latencies" nur mit T_{rr} bzw. $T_{r\bar{r}}$ in der vorliegenden Arbeit zu vergleichen. Dabei wird vorläufig vorausgesetzt, daß innerhalb des Richtig-Falsch-Paradigmas nach THORNDIKE oder BUCHWALD die Items, die mit r rückgemeldet wurden, denselben Status haben wie Items innerhalb einer Aufgabe unter der Korrekturbedingung: In beiden Fällen erfährt ja die Vp, welches die richtige Antwort ist. Das Wiederholen bzw. Nichtwiederholen dieser Antwort im nächsten (Prüf-)Durchgang führt dann zu "success" bzw. "error".

5. Latenzzeiten: Stand der Forschung

5.1 Bereichsabgrenzung

In Kapitel 1 war zur Begründung der allgemeinen Fragestellung relativ knapp bemerkt worden, daß es zum Problem der relativen Effizienz von r und f fast keine Literatur gäbe, die Latenzzeitmessung als ein wichtiges Instrument zur Behandlung dieses Problems anführe, und daß einige Autoren die Analyse von Zeitdaten als sehr erhellend für die zugrundeliegenden Gedächtnis- und Entscheidungsprozesse betrachteten. Weitere Ausführungen dazu wurden in Kapitel 1 vermieden, um die Darstellung des Zusammenhangs zwischen den Fragekomplexen (1) bis (3), (2') und (4) bis (5) nicht zu belasten.

Eben wegen des Fehlens problemspezifischer Literatur werden die folgenden Kommentare und auch die Abschnitte 3.2 bis 3.4 einen relativ weiten Bereich innerhalb der Latenzzeitforschung erfassen. "Relativ weit" heißt: Es werden nicht nur Arbeiten zum PAL unter der Nichtkorrektur-Bedingung betrachtet, sondern auch Überlegungen und Befunde, die sich ebenso auf die Messung einfacher Reaktionszeiten beziehen wie auch auf den Stellenwert von Latenzzeitmessungen innerhalb der "Strategien-Auswahl-Theorie" (nach RESTLE, 1962) bei Begriffsbildungsaufgaben. Der Zweck ist natürlich, Ansätze und Ideen, die sich in anderen Bereichen als nützlich für die Analyse von Zeitdaten erwiesen haben, für die Behandlung des vorliegenden Problems zu übernehmen, eventuell nach entsprechender Modifikation. - Der erwähnte relativ weite Bereich sei jetzt etwas erläutert.

Die älteste Tradition hat die Messung von Wahlreaktionszeiten. Hierzu gibt es auch die meisten Daten und die am sorgfältigsten ausgearbeiteten Modelle (z.B. WOODWORTH u. SCHLOSSBERG, 1954; LAMING, 1968; SMITH, 1968; STERNBERG, 1969a,b).

Die Entscheidungszeit in psychophysischen Diskriminationsexperimenten wird in neuerer Zeit besonders von AUDLEY (1960),

AUDLEY u. MERCER (1968), PIKE (1968) und ERICKSON (1970) behandelt.

Hierzu gehören auch Skalierungstechniken, die die Reaktionszeit z.B. als Maß für die Ähnlichkeit von Reizen benutzen (z.B. WOODWORTH u. SCHLOSBERG, 1954, S. 261; SIXTL, 1967, S. 246; COOMBS, 1964, S. 526). SIXTL (S. 249) kommt allerdings zu dem Schluß, daß die Entscheidungszeit kein zuverlässiges Maß für subjektive Unterschiede ergibt. Diesbezügliche Modellkonstruktionen (z.B. AUDLEY, 1960) scheinen ihm noch verfrüht.

Die PAL-Experimente, in denen Latenzzeiten gemessen wurden, sind meistens unter der Korrektur-Bedingung durchgeführt worden (z.B. HALL, 1969). Auf die wenigen Arbeiten, die eine Nichtkorrektur-Prozedur mit Zeitmessung behandeln, wird noch unter 3.2 näher einzugehen sein.

Eine Gruppe innerhalb des PAL-Bereichs sei besonders herausgestellt (WILLIAMS, 1962; EIMAS u. ZEAMAN, 1963; EIMAS, 1964; PETERSON, 1965; THEIOS, 1965; s. auch ADAMS, 1967, S. 298). Diese Autoren diskutieren die Möglichkeit, daß Latenzzeit als kontinuierliches Maß diskreten Häufigkeitsdaten hinsichtlich der Empfindlichkeit für S-R-Verbindungsstärken überlegen sei. Die Zeitmessung spielt hier also eine theoretisch wichtige Rolle in der Alles-oder-Nichts-Kontroverse (PETERSON, 1965, und THEIOS, 1965, kommen übrigens aufgrund derselben Daten zu konträren Schlußfolgerungen).

Ebenso finden auch WAUGH (1970) und WOLLEN (1970) beim Problem der assoziativen Symmetrie ("forward and backward associations"), daß Latenzzeiten empfindlicher sind als Häufigkeiten.

SCHEIRER (1971) diskutiert eingehend das Problem, inwieweit Latenzzeiten bei korrekten Antworten und Fehlerhäufigkeiten dieselben Aspekte bei Gedächtnisprozessen messen. Empirisch findet er einen nur geringen Zusammenhang zwischen diesen beiden Maßen.

Einen ebenfalls recht hohen Stellenwert hat die Latenzzeitmessung in Begriffsbildungsexperimenten (z.B. ERICKSON et al.,

1966; ERICKSON u. ZAJKOWSKY, 1967; TRABASSO u. BOWER, 1968; LEVINE, 1969; DENNEY, 1970).

Schließlich dient die Latenzzeit als Hauptmaß für die Lernleistung in Versuchen zum Vermeidungs- und Fluchtlernen (z.B. THEIOS, 1971, S. 299).

Bei der Abgrenzung des zu betrachtenden Bereichs wurde bisher unterschiedslos von Reaktions-, Entscheidungs- und Latenzzeit gesprochen. Gemeint ist immer das Zeitintervall zwischen Darbietung des Reizes S und dem Einsetzen der Reaktion R, unabhängig vom Aufgabentyp. Für diese Arbeit genügt es vorerst, sich das gesamte Zeitintervall aus drei Komponenten zusammengesetzt zu denken (nach STONE, 1960): $T = T_i + T_d + T_m$. Dabei ist T_i die Reizerkennungszeit (input time), T_d die eigentliche Entscheidungszeit (decision time), T_m die Zeit, die für die motorische Beantwortung bis zu ihrer Registrierung gebraucht wird. Differenziertere Aufspaltungen in Komponenten findet man z.B. bei CHRISTIE u. LUCE (1956), SMITH (1968), STERNBERG (1969a,b), FALMAGNE u. THEIOS (1969).

Die experimentell gefundenen Latenzzeiten enthalten natürlich im allgemeinen auch T_i und T_m , die theoretischen Überlegungen aber beziehen sich meistens auf T_d . In dieser Arbeit ist dabei T_d wiederum aufzuspalten in Gedächtnissuchzeiten und die eigentliche Entscheidungszeit für eine spezielle Reaktion. Diese Aufteilung entspricht im BM den Erinnerungsaxiomen A1, A2 und dem Entscheidungsaxiom A3.

Die Einbeziehung von Wahlreaktionszeiten, Beurteilungszeiten bei psychophysischen Diskriminationen und Latenzzeiten bei Begriffsbildungsaufgaben im Zusammenhang mit dem Problem von Latenzzeiten beim PAL soll noch weiter begründet werden. Die eben angeführten Aufgabentypen sind nicht so scharf voneinander getrennt, wie es vielleicht aufgrund der unterschiedlichen Ausdrucksweise und der Trennung in der Literatur erscheint.

Man betrachte z.B. ein Wahlreaktionszeitexperiment und eine PAL-Aufgabe, beide mit k Reaktionsalternativen: Ist die PAL-

Aufgabe hinreichend gut gelernt, reduziert sich die Latenzzeit im wesentlichen auf eine Wahlreaktionszeit (z.B. SMITH, 1968, S. 77f.).

In psychophysischen Diskriminationsaufgaben sind bei leichten Diskriminationen die Beurteilungszeiten eher mit Wahlreaktionszeiten, bei schweren Diskriminationen eher mit Latenzzeiten beim PAL vergleichbar. Die Vergleichbarkeit bezieht sich speziell auf die Differenz der Zeiten bei richtigen und falschen Antworten (PIKE, 1971; s. auch LAMING, 1968, S. 11).

Auf die fließenden Übergänge zwischen psychophysischen Diskriminierungen im Tierexperiment und Diskriminationslernen und Begriffsbildungsaufgaben im Humanbereich weisen WOODWORTH u. SCHLOSBERG (1954, S. 608f.) hin (s. auch BERGIUS, 1971, S. 65).

Die enge Verwandtschaft zwischen PAL und Begriffsbildung betonen RESTLE u. GREENO (1970, S. 101f.).

Auch wenn man vom BM ausgeht, eventuell von einer modifizierten Form, scheint es angebracht, von vornherein verschiedene "Antwortzeittypen" in die Analyse miteinzubeziehen. Je nachdem über welchen "Weg" in den B-Diagrammen eine Antwort der Vp zustandekommt, ist die zugehörige Latenzzeit eher als einfache/disjunktive Reaktionszeit oder eher als Diskriminationszeit zu betrachten. Erinnert die Vp sich z.B. sicher an R und O, speziell $O=r$, so kann sie sehr schnell reagieren; kann die Vp sich nur sehr schwach an ihre im Lerndurchgang gegebene Reaktion R erinnern, so steht sie praktisch vor der schwierigen Diskriminierungsaufgabe, aus den verschiedenen S-R-Gedächtnisspuren die relativ stärkere herauszufinden.

Die Hinweise auf die engen Beziehungen zwischen den verschiedenen Aufgabentypen sollen natürlich nicht zur Unaufmerksamkeit verleiten. Es muß in jedem einzelnen Fall überprüft werden, inwieweit ein spezifischer Befund in einem Aufgabentyp sich auf einen anderen übertragen läßt.

Zusätzlich zu den bisher unterschiedenen Experimentalsituationen ist noch darauf hinzuweisen, daß viele Berichte über

Latenzzeitmessungen sich auf Untersuchungen zum Kurzzeitgedächtnis beziehen. SHIFFRIN (1970, S. 381) bemerkt dazu, daß in diesen Fällen der Suchbereich besser zu kontrollieren und die Latenzzeit ein präziseres Maß für die Analyse der Suchprozesse darstellt; beim Wiederfinden von Gedächtnisinhalten aus dem Langzeitspeicher ist dagegen mit im ganzen längeren Zeiten zu rechnen, in denen viele unkontrollierbare Prozesse innerhalb der Vp ablaufen können (s. dazu auch auf S. 48 die kritische Anmerkung von SIXTL).

In vorliegender Arbeit sind demnach starke Schwankungen bei den Zeitdaten zu erwarten.

3.2 Experimentelle Befunde zu Fehler-Latenzzeiten

Wie in 2.5 besprochen, sind für diese Arbeit die mittleren Latenzzeiten (Mediane) T_{rr} , $T_{rr'}$, T_{ff} , $T_{ff'}$, T_{bb} und $T_{bb'}$ bzw. die Beziehungen (Rangordnungen) entsprechend Fragestellung (2) zwischen ihnen von Interesse. Daß bei Gültigkeit des BM und mit Hilfe eines darauf aufbauenden Zeitmodells präzisere Vorhersagen möglich sein können (s. Fragestellung (2')), ist in Kapitel 1 gesagt worden; Latenzzeitmodelle werden in 3.3 ausführlicher behandelt werden.

Zunächst aber sollen im folgenden Arbeiten angeführt werden, in denen die Differenz zwischen den mittleren Zeiten bei richtigen und falschen Antworten (correct and incorrect latencies, success and error latencies: im weiteren kurz T_s und T_e) diskutiert wird.

Die Struktur der meisten Experimentalsituationen ist so, daß sich die dort gefundenen Ergebnisse nur auf die Zeitmediane T_{rr} und $T_{rr'}$, wenn überhaupt, beziehen lassen. In 2.5 war dieser Punkt nur in bezug auf PAL unter der Korrektur-Prozedur erläutert worden. Aber auch z.B. bei einem Wahlreaktionszeitexperiment handelt es sich gewissermaßen um eine "Korrektur-Prozedur", da die Vp aus der Instruktion weiß, welche Reaktion zu welchem Reiz die richtige ist. Dieses gilt

auch für psychophysische Diskriminationsaufgaben. - Bei Begriffsbildungsaufgaben ist die Sachlage nicht so eindeutig. Trägt man nun Befunde zu Erfolgs- und Fehlerlatenzzeiten zusammen, ergibt sich zunächst ein uneinheitliches Bild; teils sind Erfolgszeiten kürzer als Fehlerzeiten, teils ist es umgekehrt. Es zeigt sich aber eine gewisse Systematik, wenn man die Zeiten nach den Aufgabentypen aus 3.1 ordnet.

Bei Wahlreaktionszeitexperimenten sind die Ergebnisse nach der Anzahl k der Reaktionsalternativen zu trennen. Für $k=2$ -Experimente ergeben sich übereinstimmend für falsche Reaktionen kürzere Zeiten als für richtige Reaktionen, kurz: $T_e < T_s$ (z.B. WOODWORTH u. SCHLOSBERG, 1954, S. 34; SCHOUTEN u. BEKKER, 1967; HALE, 1968; LAMING, 1968; YELLOT, 1971). Für $k > 2$ -Experimente findet LAMING (1968, S. 80), daß $T_e > T_s$ ist. Er erklärt dieses damit, daß bei anwachsendem k die Reizdiskrimination schwieriger wird. Nach FITTS (1966) ergeben sich bei einem $k=4$ -Versuch keine Unterschiede, also $T_e \approx T_s$. (Er diskutiert im übrigen die Abhängigkeit solcher Befunde von der Instruktion, ob sie mehr Wert auf Schnelligkeit oder Korrektheit der Reaktion legt: "speed vs. accuracy".)

In psychophysischen Diskriminationsexperimenten ($k=2$) ergibt sich folgendes Bild: Bei leichten Diskriminierungen ist $T_e < T_s$, bei schweren Diskriminationen ist $T_e > T_s$ (PIKE, 1968, 1971; AUDLEY u. MERCER, 1968). PIKE (1968) weist dabei darauf hin, daß im letzteren Fall durch "speed"-Instruktionen ein Umkippen erfolgen kann, also $T_e < T_s$. PIKE (1971) betont die Ähnlichkeit von leichten Diskriminationsaufgaben mit Wahlreaktionsversuchen (für $k=2$). Schwerere Diskriminationsaufgaben dagegen haben mehr Ähnlichkeit mit $k > 2$ -Wahlreaktionen, jedenfalls hinsichtlich der Entscheidungsschwierigkeiten.

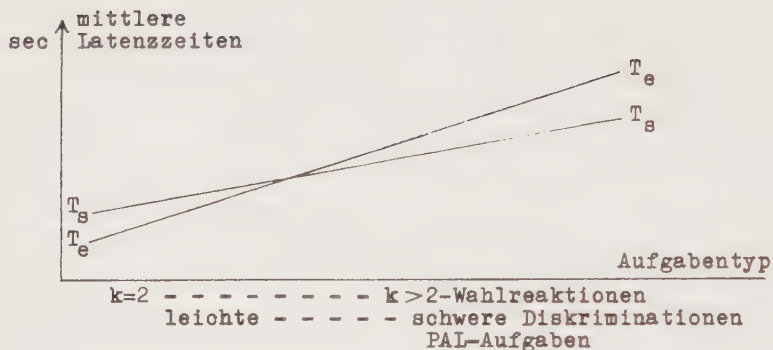
Ein Diskriminationsexperiment unter verschiedenen Bestrafungsbedingungen (leises, mittleres, lautes Geräusch) beschreibt ERICKSON (1970): Hier ergibt sich ein U-förmiger Zusammenhang zwischen sowohl Unterscheidungsleistung als auch Latenzzeiten und dem Bestrafungsniveau; bei mittlerem Geräusch resultieren die größte Leistung und die längsten Latenzzeiten.

Bei den folgenden PAL-Experimenten werden zunächst nur Arbeiten zitiert, in denen die Korrektur-Methode angewandt und in denen eine größere Anzahl von Durchgängen durchgeführt wurde (also nicht wie in den BUCHWALDschen Experimenten mit je einem Lerndurchgang und einem Prüfdurchgang). Um die Latenzzeitdaten aus diesen Experimenten mit T_{rr} und $T_{r\bar{r}}$ in Beziehung setzen zu können, dürfen daher nicht die gesamten Latenzzeitkurven ("conditionalized on errors and successes", s. SUPPES et al., 1966, S. 118) betrachtet werden, sondern nur die Werte T_s und T_e , die man über dem zweiten Durchgang ($n=2$) abliest. Der erste Durchgang entspricht dem Lerndurchgang bezüglich der r-Wörter in einem BUCHWALD-Experiment, der zweite Durchgang dem Prüfdurchgang. Es sind also T_s und T_e für $n=2$ mit T_{rr} und $T_{r\bar{r}}$ vergleichbar.

Es ergeben sich folgende Befunde: $T_s < T_e$ ($n=2$) bei MILLWARD (1964a; $k=2$), KELLER et al. (1965; $k=2$), SUPPES et al. (1966; $k=3$), HALL (1969; $k=4,8$), HALL u. WENDEROTH (1972; $k=4,8$). Auch aus den etwas anders aufgebauten Experimenten von EIMAS u. ZEAMAN (1963), EIMAS (1964), NORMAN u. WICKELGREN (1969), WOLLEN (1970), WEARING (1970, 1971) erhält man: $T_s < T_e$.

Nicht so klar sind die Ergebnisse bei KELLER et al. (1967; $k=3$; $T_s < T_e$ u n d $T_s > T_e$), bei HALL (1969; $k=2$) und HALL u. WENDEROTH (1972; $k=2$). Bei den beiden letzten Arbeiten sind die entsprechenden Diagramme nicht ganz eindeutig interpretierbar: Es scheint sich eine leichte Tendenz für $T_e < T_s$ abzuzeichnen. Kompliziert wird das Ergebnis auch dadurch, daß in der Arbeit von HALL u. WENDEROTH (1972) unter der $k=2$ -Bedingung spezielle Strategien untersucht werden. Abgesehen von diesen Ausnahmen ergibt sich eine starke empirische Begründung für die später (s. Kapitel 4) abzuleitende Hypothese: $T_{rr} < T_{r\bar{r}}$.

Die bisherigen Befunde zum Verhältnis von T_s und T_e in Abhängigkeit vom Aufgabentyp lassen sich schematisch darstellen. (Die mittleren Zeiten bei Wahlreaktionen sind natürlich kürzer als z.B. bei schwierigen Diskriminationen.)



Diese Skizze sagt nur etwas über zentrale Tendenzen aus; ebenso wie bei Wahlreaktionszeiten ($k=2$: $T_e < T_s$) kann es selbstverständlich auch bei einzelnen PA-Items zu übereilten Reaktionen kommen, d.h. zu Fehlerzeiten, die kürzer als die Erfolgszeiten sind.

Die bisher angeführten Befunde bezogen sich auf das Verhältnis von T_s und T_e (bzw. T_{rr} und $T_{r\bar{r}}$). Kann man aus ihnen auch Aussagen bezüglich T_{ff} und $T_{f\bar{f}}$ entnehmen?

Beim PAL sind $k=2$ -Aufgaben unter der Nichtkorrektur-Bedingung vom reinen Informationsstandpunkt aus betrachtet, äquivalent zu entsprechenden Aufgaben unter der Korrektur-Bedingung. Sowohl nach r als auch nach f ist die richtige Reaktion für den nächsten Durchgang eindeutig bestimmt. Die Erweiterung der empirisch gefundenen Aussage: $T_s < T_e$, auf $T_{f\bar{f}}$ (Erfolgszeit) und T_{ff} (Fehlerzeit) ist aber trotzdem problematisch. Das möglicherweise stattfindende "Umschalten" bei $f\bar{f}$ -Antworten kann eventuell das Zeitverhältnis umkehren. D.h., gleiche Informationsverhältnisse müssen nicht unbedingt gleiche Zeitverhältnisse implizieren (s. auch ähnliche Überlegungen bei WAUGH, 1970; WOLLEN, 1970).

In den drei folgenden Arbeiten wurden Latenzzeitmessungen unter einer "noncorrection"-Prozedur durchgeführt.

Für die erste dieser Arbeiten (KELLER et al., 1965) muß aber

sofort eine Einschränkung gemacht werden. Die Verstärkungen nach der Reaktion der Vp bestehen nicht in den Rückmeldungen r bzw. f, sondern in "Punkten" (diese stehen in Beziehung zur späteren Bezahlung der Vp). Für jedes Reizwort in der Lernliste ist den beiden Reaktionsalternativen ein Punktepaar zugeordnet, z.B. 1 und 8 Punkte, 2 und 4 Punkte usw. (Es wurden sämtliche Kombinationsmöglichkeiten von Punktwerten benutzt, die sich aus den Werten 1 bis 8 bilden lassen.) Die Reaktionsalternative mit dem höheren Punktwert ist die "richtige" Reaktion.

Unter der "noncorrection"-Prozedur erfährt die Vp nur den Punktwert, der ihrer gegebenen Reaktion zugeordnet ist. Bei 8 (bzw. 1) ist die Vp dann sicher, daß ihre Antwort richtig (bzw. falsch) war; bei 4 dagegen weiß die Vp nicht, ob die 4 zu einem Punktepaar wie 2-4 gehört oder zu einem Punktepaar wie 4-7. Im ersteren Falle würde 4 r bedeuten, im letzteren Falle f.

Die "noncorrection"-Prozedur resultiert hier also in einer Unsicherheit, ob die Rückmeldung, z.B. 4, r oder f bedeutet, während unter der bisher beschriebenen Nichtkorrektur-Bedingung bei $k > 2$ und f-Rückmeldung die Unsicherheit darin liegt, welche der übrigen Antwortalternativen nun die richtige ist. Die Vergleichbarkeit der beiden, zwar gleich benannten, Prozeduren ist also problematisch.

Nichtsdestoweniger kommen KELLER et al. zu einem Resultat, daß sich in Beziehung zu dem Verhältnis T_{rr} und T_{ff} setzen läßt: Wenn der Punktwert relativ hoch ist, z.B. 7, akzeptiert die Vp die dazugehörige Reaktion als richtige und antwortet im nächsten Durchgang relativ schnell (s. T_{rr}). Wenn der Punktwert relativ niedrig ist, z.B. 2, neigt die Vp dazu, die dazugehörige Reaktion zurückzuweisen und an deren Stelle die andere zu probieren, ein Umschaltprozeß, der Zeit braucht (s. T_{ff}).

Die beiden anderen Untersuchungen unter Nichtkorrektur-Bedingungen von BERGHÄUSER u. SCHRÖDER (1972) und HEIDE (1974)

sind den Experimenten in vorliegender Arbeit einigermaßen ähnlich. Es ergaben sich keine interpretierbaren Abhängigkeiten der mittleren Latenzzeiten von r , f und b . Es ist möglich, daß hier versuchstechnische Gründe, wie z.B. relativ schnelle Versuchsdurchführung oder die Vpn irritierende blinkende Lämpchen, die Hervorbringung eventuell vorhandener Effekte verhinderten.

Es seien nun einige Arbeiten zur Begriffsbildung besprochen. ERICKSON et al. (1966) und ERICKSON u. ZAJKOWSKY (1967) gehen von den "Strategien- bzw. Dimensionen-Auswahl"-Modellen von RESTLE (1962) bzw. BOWER u. TRABASSO (1964) aus (s. auch LEVINE, 1969).

Es wird dort angenommen, daß der Antwort einer Vp in einem bestimmten Durchgang ein Hypothesen-"pool" zugrundeliegt. Die Erweiterung auf Latenzzeiten erfolgt durch eine einfache Annahme: Latenzzeiten sind eine monotone Funktion von H , der Größe des Hypothesen-"pool". Daraus folgt u.a. die Voraussage: Latenzzeiten sind länger, wenn sie f -Rückmeldungen als wenn sie r -Rückmeldungen folgen. Im ersteren Fall ist $H=2^D$ (D binäre Dimensionen vorausgesetzt), im letzteren Fall $H=1$. Obige Voraussage wird durch die Daten bestätigt.

Diese Überlegungen lassen sich angenähert auf $k > 2$ -Situationen übertragen. Falls r und f (und R) im Prüfdurchgang korrekt erinnert werden, antwortet die Vp nach r aus einem "pool" bestehend aus einer Antwort, nämlich der richtigen, nach f aus einem "pool" bestehend aus $k-1$ Antworten. Daraus folgt dann $T_{rr} < T_{ff}$.

In einer weiteren Arbeit zur Begriffsbildung (DENNEY, 1970; s. auch GREENWALD, 1970) wurden außer den Latenzzeiten auch die Zeiten gemessen, die die Vp sich nach der Rückmeldung nahm, bevor sie sich das nächste Reizmuster projizieren ließ. Diese Zeiten waren nach f länger als nach r . Die Informationsverarbeitung nach f ist also aufwendiger als nach r .

Zum Abschluß sei auf die allgemeine Abhängigkeit der Antwortzeiten von der Anzahl der Reize und der möglichen Reaktions-

alternativen hingewiesen (s. z.B. für den Bereich der Wahlreaktionszeiten SMITH, 1968, S. 80f.): Im allgemeinen wächst die Reaktionszeit mit komplexer werdender Wahlsituation an, z.B. mit ansteigender Antwortalternativenzahl k .

Mit welchen Zeitdifferenzen in Abhängigkeit von k hat man zu rechnen?

Für reine Wahlreaktionszeiten wurde gefunden (WOODWORTH u. SCHLOSBERG, 1954, S. 33): $T_{k=2} = 0,316$ sec; $T_{k=4} = 0,434$ sec. Es ergibt sich als Differenz ca. 0,12 sec. - Die Werte für $k=2$ und $k=4$ wurden zitiert, weil diese Alternativenanzahlen in der vorliegenden Arbeit benutzt werden.

Aus dem PAL-Bereich können die Arbeiten von HALL (1969, S. 124) und HALL u. WENDEROTH (1972, S. 33) herangezogen werden. Hier sind die interessierenden Differenzen ($k=2$ und $k=4$) wieder über dem zweiten Durchgang ($n=2$) abzulesen. Die absoluten Zeiten bewegen sich im Bereich von 1,1 bis 1,5 sec. Die Differenzen schwanken zwischen 0,1 und 0,25 sec ; sie sind dabei für die Fehlerzeiten etwas größer als für die Erfolgszeiten. (Die Instruktion hob die Erwünschtheit schnellen Antwortens hervor.)

Diese Abhängigkeit von k kann eventuell abgeschwächt werden, wenn die Instruktion mehr die Korrektheit als die Geschwindigkeit des Antwortens betont. In diesem Fall kann der Effekt der subjektiven (Un-)Sicherheit den der k -Abhängigkeit überdecken (z.B. WEARING, 1971). Ebenfalls im Auge behalten muß man die Möglichkeit, daß Vpn im Falle sehr schwacher Erinnerungen an ihre Reaktionen im Lerndurchgang bei $k=4$ eher resignieren und relativ schneller eine geratene Antwort abgeben als bei $k=2$ (s. dazu EDWARDS, 1965; POLLAY, 1970).

3.3 Latenzzeitmodelle

Wie in Kapitel 1 in der Vorausschau erwähnt, erwies sich das BM nicht als hinreichend geeignet, um darauf ein mathematisches Latenzzeitmodell aufzubauen. Die Fragestellung (2'b)

entfiel damit. Daher werden im folgenden Überblick die Latenzzeitmodelle nicht so ausführlich dargestellt, wie es für die Bewältigung von (2'b) notwendig gewesen wäre. Es sollen besonders die Modelle hervorgehoben werden, aus denen sich Anhaltspunkte für die Ableitung von Hypothesen in Kap. 4 entsprechend Fragestellung (2) ergeben.

Allgemeine Übersichten über Zeitmodelle, speziell für Wahlreaktionszeiten findet man bei SMITH (1968) und LAMING (1968). Die mathematische Behandlung von stochastischen Latenzzeitmechanismen stellt MCGILL (1963) ausführlich dar (s. ferner MCGILL u. GIBBON, 1965). Auch STERNBERG (1966, 1969a,b) gibt allgemeinere Methoden und Überlegungen zur Behandlung von Reaktionszeitverteilungen an, die er dann selbst auf einen Spezialfall anwendet: Die V_p muß angeben, ob ein Testsymbol in einer kurz zuvor eingeprägten Liste enthalten war oder nicht.

Relativ frühe Arbeiten zur Verbindung von Latenzzeiten mit Lernmodellen findet man bei HULL (1943, S. 344), MUELLER (1950), ESTES (1950, 1959), BUSH u. MOSTELLER (1955, S. 310f.). Das Modell von BUSH u. MOSTELLER ist besonders interessant, weil diese Autoren eine sehr enge Verflechtung des Zeitmechanismus mit ihrem Lernmodell anstreben und versuchen, jedem Parameter eine psychologische Bedeutung zu geben (s. MCGILL, 1963, S. 323). Diese enge Verflechtung ist ja auch im Zusammenhang mit dem BM in Hinblick auf Fragestellung (3) von Bedeutung.

Von CHRISTIE u. LUCE (1956) stammt eine Behandlung von Entscheidungszeiten in Wahlsituationen, die auf die B-Diagramme anwendbar erscheint. Die Vorgänge beim Entscheidungsprozeß werden durch einen "gerichteten Graphen" dargestellt, der eine Anzahl von Elementarentscheidungen beinhaltet, die schließlich zu einer Gesamtentscheidung für eine bestimmte Reaktion führen. Entsprechend setzt sich die Gesamtentscheidungszeit, abgesehen von einer konstanten Basiszeit, aus mehreren elementarerer Entscheidungszeiten zusammen. Es wird ferner postuliert, daß diese Elementarzeiten die gleiche Verteilung haben und unabhängig voneinander sind.

Dieses letztere Postulat hat den Vorteil, die Anwendung der LAPLACE-Transformation zu ermöglichen. Hierdurch wird die Behandlung einer Summe von unabhängigen Elementarzeiten einfach, da die LAPLACE-Transformierte der Dichte einer Summe von unabhängigen Zufallsvariablen gleich dem Produkt der LAPLACE-Transformierten der einzelnen Dichten ist. (Auf die sehr enge mathematische Beziehung zu den momenterzeugenden Funktionen sei hingewiesen; s. ATKINSON et al., 1965, S. 168.)

Um die mathematische Behandlung weiter zu vereinfachen, werden als spezielle Graphen rein serielle bzw. parallele Entscheidungsgraphen diskutiert (s. dazu LAMING, 1966; BOWER, 1967).

Für die Elementarzeitenverteilung wird eine Exponentialverteilung angenommen. Diese in der Literatur zur theoretischen Behandlung von Latenzzeiten sehr häufig gemachte Annahme führt im Falle des serialen Graphen zu einer Gamma-Verteilung (z.B. MCGILL, 1963; für den Fall unterschiedlicher Zeitkonstanten s. speziell MCGILL u. GIBBON, 1965). Die allgemein bevorzugte Gamma-Verteilung wird übrigens auch kritisiert (z.B. PIKE, 1968; SNODGRASS et al., 1967).

Bei der Übertragung der Vorgehensweise von CHRISTIE u. LUCE auf die B-Diagramme ist allerdings Vorsicht geboten. Die Darstellung auf S. 31 legt zwar die Vorstellung von "gerichteten Graphen" nahe, es ist aber durchaus nicht sicher, ob die B-Diagramme auch eine zeitliche Ordnung implizieren, z.B. in dem Sinne, daß erst ein Suchprozeß und eine Entscheidung bezüglich der S-R-Verbindung und dann entsprechend für die S-O-Verbindung stattfinden. Die Reihenfolge kann durchaus auch umgekehrt sein, bzw. die Prozesse bezüglich der S-R- und der S-O-Verbindungen können parallel ablaufen. Die Entscheidung für die endgültige Antwort erfolgt wohl meistens, bis auf "überstürzte" Reaktionen, nach den Such- und Entscheidungsprozessen hinsichtlich der S-R- und S-O-Erinnerungen.

Es soll jetzt speziell der " $c \cdot (1-u)$ -Weg" aus den beiden B-Diagrammen für r und f näher betrachtet werden, da er später

experimentell isoliert werden soll und für die Ableitung der Hypothesen in Kapitel 4 eine wichtige Rolle spielt. Falls dieser "Weg" als ein serialer gerichteter Graph angenommen werden kann und falls die Unabhängigkeitsannahme U₃ BUCHWALDs bezüglich der Erinnerungen an R und O gültig ist (s. S. 34), scheint die Anwendung des LAPLACE-Formalismus möglich. Dabei ist noch vorausgesetzt, was nicht selbstverständlich ist, daß dann auch die Suchzeiten bezüglich R und O unabhängig voneinander sind.

Falls sich aber U₃ als falsch erweisen sollte, sind sicherlich auch die zugehörigen Prozeßzeiten voneinander abhängig, und die bekannten Ableitungen, die z.B. zur, eventuell verallgemeinerten, Gamma-Verteilung führen, werden unanwendbar. Dieses bleibt richtig, auch wenn sich die Zeiten noch additiv zusammensetzen sollten. (Zur Additivität ohne gleichzeitige stochastische Unabhängigkeit s. STERNBERG, 1969a).

Eine große Gruppe von Ansätzen zur Behandlung von Latenzzeiten in Wahlsituationen läßt sich, wie folgt, grob umreißen (eine differenziertere Darstellung gibt LAMING, 1968; s. auch ATKINSON et al., 1965; AUDLEY u. PIKE, 1965).

In einer Wahlsituation passieren vor der endgültigen "offenen" (overt) Entscheidung für eine bestimmte Reaktion aus mehreren möglichen eine Reihe von hypothetisch angenommenen "verdeckten" (covert) Ereignissen, die den verschiedenen "offenen" Reaktionen zugeordnet sind. (Es werden auch neutrale Ereignisse angenommen, die zu keiner Reaktion gehören, aber die Entscheidungszeit verlängern.)

Solche "verdeckten" Ereignisse werden unterschiedlich benannt und interpretiert, z.B. "Orientierungs- und Annäherungsreaktionen" (BOWER, 1959), "implizite Antworten" (AUDLEY, 1960), "Reizelemente" (LaBERGE, 1962), "Informationszuwächse" gleich "Aufteilung eines Stroms von Informationen in eine Reihe von diskreten Zufallsvariablen" (STONE, 1960; LAMING, 1968; s. auch EDWARDS, 1965; FITTS, 1966).

Diese "verdeckten" Ereignisse werden nun, getrennt für die verschiedenen "offenen" Reaktionen, registriert, eventuell

auch noch geeignet transformiert. Sobald nun die Ergebnisse dieser Registrierung bestimmte Schwellen (absolute oder relative) überschreiten, erfolgt die eine oder andere "offene" Reaktion.

Nach z.B. LaBERGE (1962; "recruitment"- oder "accumulator-model") werden in einer $k=2$ -Wahlaufgabe bei Darbietung der Reizsituation so lange "Reizelemente", die mit der einen oder anderen Reaktion assoziiert sind, in zwei getrennten Zählern registriert, bis schließlich bei einem der beiden Zähler ein Schwellenwert ("recruitment factor" r') überschritten wird und die zugehörige offene Reaktion erfolgt. Die Antwortzeit ist also eine Funktion von dem "recruitment"-Faktor r' . - Eine interessante Erweiterung findet man bei NICKERSON (1969), der noch einen zusätzlichen Zähler einführt, der die während des Entscheidungsvorganges verfließende Zeit registriert und ebenfalls zur endgültigen Entscheidung beitragen kann.

Das Modell von LaBERGE wird von PIKE (1968, 1971) und AUDLEY u. MERCER (1968) favorisiert, da der "recruitment"-Faktor r' empfindlich auf motivationale Faktoren anspricht. Je nach den Annahmen über r' kommt man auch zu unterschiedlichen Voraussetzungen bezüglich des Verhältnisses von T_s und T_e . (ERICKSON, 1970, findet allerdings in seinem Experiment Übereinstimmung des "recruitment"-Modells nur mit den Wahlhäufigkeiten, nicht mit den Latenzzeiten.)

Einen besonderen Effekt sagt das Modell von LaBERGE (1962, S. 389f.) bezüglich der Differenz $T_e - T_s$ in Abhängigkeit von dem "recruitment"-Faktor r' voraus: Wird z.B. durch Instruktion r' vergrößert, speziell durch Betonung von Genauigkeit der Reaktionen, müssen also mehr "Reizelemente" registriert werden, bevor eine "offene" Reaktion stattfindet, so wachsen sowohl T_e als auch T_s für sich an, insbesondere wird aber auch die Differenz $T_e - T_s$ markanter. Auf diesen Effekt wird unter 6.3.4 Bezug genommen.

Eine andere große Gruppe von Latenzzeitmodellen ist dadurch gekennzeichnet, daß von Markov-Modellen des Lernens ausgegangen wird, die sich schon im Bereich von Häufigkeitsdaten gut bewährt haben (s. zur Einführung ATKINSON et al., 1965).

Den einzelnen, hypothetisch angenommenen Lernzuständen C_i werden unterschiedliche mittlere Antwortzeiten μ_i zugeordnet. Seltener werden auch Varianzen oder Verteilungsformen mitbehandelt. Die Voraussage für die mittlere Latenzzeit im Durchgang n ist dann das gewichtete Mittel aus den μ_i . Die Gewichte sind dabei die einzelnen Wahrscheinlichkeiten $u_{i,n}$ für die Zustände C_i im Durchgang n , die sich unabhängig von Latenzzeitbetrachtungen aus dem jeweiligen Markov-Modell berechnen lassen. Die μ_i sind als zusätzliche Parameter noch aus den Zeitdaten zu schätzen.

Eventuell werden auch schon aus gewissen Modellvorstellungen heraus Aussagen über die voraussichtliche Rangreihenfolge der μ_i abgeleitet.

Zu dieser so charakterisierten Gruppe gehören die Arbeiten von ATKINSON u. CROTHERS (1964), MYERS u. ATKINSON (1964), MILLWARD (1964a), THEIOS (1965, 1971), SUPPES et al. (1966), MYERS et al. (1967), HALL (1969), HALL u. WENDEROTH (1972). Einige dieser Untersuchungen sind auch schon unter 3.2 angeführt. Dort wurden allerdings nur die empirischen Befunde im Durchgang $n=2$ bezüglich T_s und T_e benötigt.

Für die Argumentation, die später in vorliegender Arbeit benutzt werden soll, genügt als Ausgangspunkt das "Ein-Element"-Modell (BOWER, 1961), auf dem die Arbeit von MILLWARD (1964a) basiert. Gewisse Abweichungen von den Schlußfolgerungen scheinen erst bei größerem k (z.B. $k=8$) relevant zu werden (z.B. HALL, 1969).

Das "Ein-Element"-Modell besagt auf PAL angewandt kurz: Ein Item befindet sich entweder in einem ungelernten (Rate-)Zustand C_1 oder in einem gelernten Zustand C_2 . Der Übergang vom anfänglichen Zustand C_1 zum Endzustand C_2 (in diesem Modell ist Vergessen ausgeschlossen) erfolgt mit der Wahrschein-

lichkeit c in irgendeinem Durchgang. Der Lernparameter c ist aus den Daten zu schätzen. Danach macht die V_p bei diesem Item keine Fehler mehr.

Betrachtet man den Durchgang, bei dem der letzte Fehler für dieses Item registriert wurde, so weiß man mit Sicherheit, dem Modell nach, daß sich das Item in diesem Durchgang und in allen vorherigen Durchgängen noch im ungelernten (Rate-) Zustand C_1 befunden hatte. Etwaige richtige Antworten sind nur durch Raten zustandegekommen. Es ist bei der Analyse der Daten üblich, die einzelnen Fehler-Erfolgs-Sequenzen für die verschiedenen Items so anzuordnen, daß sie mit dem jeweiligen Durchgang des letzten Fehlers übereinstimmen.

Hat man nun, wie oben beschrieben, dem Zustand C_1 die mittlere Antwortzeit μ_1 zugeordnet, so weiß man, daß die Latenzzeiten für alle Antworten im und vor dem Durchgang des letzten Fehlers aus derselben Verteilung mit dem Mittelwert μ_1 stammen. Insbesondere müssen also die mittleren Zeiten für richtige und falsche Antworten bis auf Zufallsschwankungen gleich sein. Diese Vorhersage wird im allgemeinen gut bestätigt (s. besonders deutlich in KELLER et al., 1967).

Nach dem Durchgang des letzten Fehlers geht das Item früher oder später in den (absorbierenden) gelernten Zustand C_2 über. Hier erfolgen dann die Antwortzeiten aus der Verteilung mit dem Mittelwert μ_2 . Die Daten (z.B. MILLWARD, 1964a) ergeben eindeutig: $\mu_2 < \mu_1$ (s. dazu auch die beiden aufeinander bezogenen Arbeiten von PETERSON, 1965, und THEIOS, 1965).

Diese Relation $\mu_2 < \mu_1$ kann auch theoretisch vorausgesagt werden. Die Argumentation HALLs (1969) dazu wird stark vereinfacht und an die augenblickliche Darstellung angepaßt wiedergegeben: Im ungelernten Zustand C_1 durchmustert ("scans") die V_p bei Darbietung des Reizwortes alle Antwortalternativen, ohne zu einem "Erfolg" ("match") zu kommen und rät schließlich. Im gelernten Zustand C_2 gelangt die V_p im Durchschnitt früher zu einem "match". Daraus folgt: $\mu_2 < \mu_1$.

Aus dieser Argumentationsweise ergibt sich nach HALL auch die Abhängigkeit der Zeiten von den Antwortalternativen k .

Bei größerem k dauert das Durchmustern eben länger.

Bei der Ableitung der Hypothesen in Kapitel 4 wird auf die obigen Überlegungen zum "Ein-Element"-Modell Bezug genommen. Mit aller Vorsicht werden dabei der " $c \cdot (1-u)$ -Weg" mit dem gelernten Zustand C_2 und die "restlichen" Wege mit dem ungelerten Zustand C_1 in Beziehung gesetzt. Insbesondere wird angenommen werden, daß alle Zeiten, die über die "restlichen" Wege zustandekommen, sich nicht wesentlich unterscheiden, ob es sich dabei nun um richtige oder falsche Antworten handelt.

Die oben angemahnte Vorsicht soll noch etwas erläutert werden. In den Markov-Modellen hat die Verstärkung (reinforcement) die Aufgabe, den Übergang eines Items z.B. vom ungelerten in den gelernten Zustand zu bewirken. Die Verstärkung selbst taucht dann aber in den Zuständen nicht mehr auf. Anders ist es beim BM: Hier erscheint die Verstärkung, also die Rückmeldung, explizit "längs" des " $c \cdot (1-u)$ -Weges", eben als S-O-Erinnerung. Es ist also nicht unproblematisch, "Wege" mit "Zuständen" zu vergleichen.

Dieser Punkt wird so betont, weil der Vf. dieser Arbeit zeitweilig versucht war, die S-R- bzw. S-O-Assoziationen als "Reizelemente" im Sinne der "Reiz-Stichproben-Theorie" (s. ATKINSON u. ESTES, 1963) zu interpretieren. Wegen der ausgesprochenen Flexibilität dieser Rahmentheorie bezüglich der Definition von "Reizelementen" schien diese Betrachtungsweise nicht ganz abwegig.

Aus der obigen Einführung in das "Ein-Element"-Modell in der auf Latenzzeiten erweiterten Form ist zu entnehmen, daß bei den richtigen Antworten unmittelbar nach dem Durchgang des letzten Fehlers nicht entscheidbar ist, ob sie schon aus dem Lernzustand C_2 erfolgen oder noch aus C_1 kommen, aber per Zufall richtig sind. Für diesen Bereich ist die beobachtete Latenzzeitverteilung ein Wahrscheinlichkeitsgemisch aus den μ_1 - und μ_2 -Verteilungen.

Die Darstellung einer Verteilung aus mehreren Verteilungskomponenten wird oft angewandt (z.B. FALMAGNE u. THEIOS,

1969). Es liegt nahe, bei Gültigkeit des BM und stabilen Parameterschätzungen $\hat{c}$, $\hat{d}$ und $\hat{g}$ die beobachteten Zeitverteilungen als Wahrscheinlichkeitsgemische der einzelnen "Weg"-Verteilungen zu analysieren; dabei wären zusätzlich noch einige "Elementarzeiten"-Parameter bezüglich der Gedächtnis- und Entscheidungsprozesse aus den Zeitdaten zu schätzen.

Auf jeden Fall wird die Grundidee eines Gemisches übernommen werden, und zwar in folgender Form: Die Gesamtverteilung besteht aus einer relativ "langsamen" Grundkomponente, der unterschiedliche Anteile von "schnellen" Zeiten zugemischt sind.

3.4 Latenzzeiten und Konfidenzurteile

Bisher ist die Abhängigkeit der Latenzzeiten von Aufgabentyp, Antwortalternativenzahl k , Korrektheit der Antworten und Instruktion angeschnitten worden. Auf einen Faktor, der vermutlich starken Einfluß auf die Dauer der Latenzzeiten haben wird, soll kurz hingewiesen werden. Es handelt sich dabei um die subjektive Sicherheit (confidence), mit der Vpn ihre Reaktionen abgeben. Dabei kann, z.B. im Rahmen des BM, die subjektive Sicherheit von mehreren Komponenten herrühren, die sich z.B. auf die Erinnerungen an R bzw. O oder auf die anschließenden Entscheidungen beziehen.

Im allgemeinen findet man experimentell, daß mit wachsender Sicherheit die Antwortzeiten kürzer werden und umgekehrt (s. z.B. WOODWORTH u. SCHLOSBERG, 1954, S. 263; WEARING, 1970). Dabei muß hohe Sicherheit nicht unbedingt Korrektheit der Antwort bedeuten. Es kann auch eine falsche Antwort mit hoher subjektiver Sicherheit, also relativ schnell gegeben werden (s. WEARING, 1971).

Theoretische Modelle müssen diesen Befunden Rechnung tragen (z.B. AUDLEY, 1960; BOWER, 1967, S. 239 u. 292). AUDLEY macht den Ansatz: Latenzzeit $\sim 1/\text{Sicherheit}$. Nach ATKINSON et al. (1965, S. 163) ist dieses eine ad-hoc-Annahme und im

wesentlichen durch die Alltagsvorstellung von dem Zusammenhang zwischen Antwortzeit und subjektiver Sicherheit gerechtfertigt.

Von der Übertragung des signal-entdeckungs-theoretischen Ansatzes auf Wiedererkennungstests her (s. EGAN, 1958) argumentieren NORMAN u. WICKELGREN (1969). Sie erweitern diese Überlegungen auf Mehrfachwahlprüfungen mit der zusätzlichen Annahme, daß bei wenigen und gut bekannten Antwortalternativen die Wiedererkennungs- und die Mehrfachwahlprüfung sich sehr ähnlich sind (s. dazu auch ATKINSON u. SHIFFRIN, 1968, S. 122). Sie kommen ebenfalls zu der Folgerung: Schnelle (langsame) Antworten, ob richtig oder falsch, sind mit hoher (niedriger) subjektiver Sicherheit verbunden (s. auch MURDOCK u. DUFTY, 1972; diese Autoren stellen allerdings fest, daß die theoretischen Vorhersagen nur für die mittleren Antwortzeiten stimmen, nicht aber für die Varianzen).

WEARING (1971) findet komplexere Zusammenhänge: Die Abhängigkeit der Latenzzeiten von der subjektiven Sicherheit ist unterschiedlich für richtige und falsche Antworten. Die Zeiten hängen auch noch vom Lernmaterial ab, z.B. vom Assoziationswert der Items (ARCHER, 1960) oder von der Möglichkeit, zu einzelnen Items "Eselsbrücken" (NLMs = natural language mediators) bilden zu können (MONTAGUE u. KIESS, 1968). Die im wesentlichen reziproke Beziehung zwischen Latenzzeit und subjektiver Sicherheit bleibt aber unbestritten.

Es ergibt sich eine Folgerung für das nächste Kapitel (Ableitung von Einzelhypothesen bezüglich der sechs Zeitmediane T_{rr} , $T_{r\bar{r}}$, T_{ff} , $T_{f\bar{f}}$, T_{bb} und $T_{b\bar{b}}$). Dort wird argumentiert werden, daß einige Mediane deswegen kürzer sind als andere, weil zu ihnen ein bestimmter Anteil von subjektiv sicheren, also schnellen Antworten beiträgt.

4. Latenzzeiten: Hypothesen

Die im folgenden abgeleiteten Einzelhypothesen bezüglich des Verhältnisses je zweier Mediane zueinander stellen eine Differenzierung bzw. Präzisierung der allgemeineren Fragestellungen (1) und (2) dar (s. S. 3).

Die Ableitungen sind unabhängig von der strikten Gültigkeit des BM (1969). Insbesondere brauchen die Unabhängigkeitsforderungen U1 bis U3 (s. S. 34) nicht vorausgesetzt zu werden. Beibehalten werden soll aber die gedächtnispsychologische Interpretation des "Verstärkungs"-Vorgangs nach BUCHWALD: Es werden also weiterhin zwei Gedächtnisprozesse, ob abhängig oder unabhängig voneinander, bezüglich der Reaktionen R und der Rückmeldungen O aus dem Lerndurchgang vorausgesetzt. Wie schon auf S. 2 erläutert, wird die Gültigkeit der letzten Voraussetzung allgemein durch die kognitiv gehaltene Versuchssituation nahegelegt. In einigen Experimentalgruppen wird sie noch besonders intensiv durch die spezielle Prüftechnik durchgesetzt ("Verbalisieren").

Sollte sich das BM als gültig erweisen und kann auf ihm basierend ein Latenzzeitmodell erfolgreich konstruiert werden, so werden die hier abgeleiteten Hypothesen nicht hinfällig, sondern müßten sich sogar in präziserer Form aus diesem Modell ergeben. Mit "präziser" ist gemeint: Statt Größer-Kleiner-Beziehungen zwischen je zwei Medianen sollten sich Gesamtreihenfolgen der sechs Mediane bzw. auch numerische Differenzen zwischen ihnen voraussagen lassen.

Auf jeden Fall wird jetzt versucht werden, in möglichst einheitlicher Vorgehensweise aus den B-Diagrammen bzw. einer modifizierten Form dieser Diagramme Einzelhypothesen bezüglich je zweier Mediane abzuleiten.

Die "einheitliche" Vorgehensweise sei betont, d.h. sie soll angenähert dem Ableitungsverfahren bei einem mathematischen Modell entsprechen. Dort werden aus einigen wenigen Axiomen zahlreiche Vorhersagen abgeleitet.

Die Absicht ist, einen zu starken Bruch in der Behandlung der Fragestellungen (2) bzw. (2'b) zu vermeiden. Daher werden im

"Stand der Forschung", besonders in 3.2 und 3.3, angegebene Befunde hauptsächlich zur Unterstützung der abgeleiteten Hypothesen herangezogen, weniger aber zur Ableitung selbst, zumal bei diesen Befunden oft nicht die direkte Vergleichbarkeit mit den hier durchgeführten Experimenten vorausgesetzt werden kann. Die Begründungen der Hypothesen aus den B-Diagrammen und aus der Literatur gehen aber natürlich fließend ineinander über.

Wie in 3.1 bemerkt, wird im weiteren nur T_d analysiert werden; $T_i + T_m = \text{konst.}$ wird vorausgesetzt.

Der Anschauung halber werden die B-Diagramme (s. S. 31) nochmals aufgeführt. Die Modifikationen in der Darstellungsweise werden dann anschließend erläutert (s. auch in 6.3.1 zum modifizierten BM).

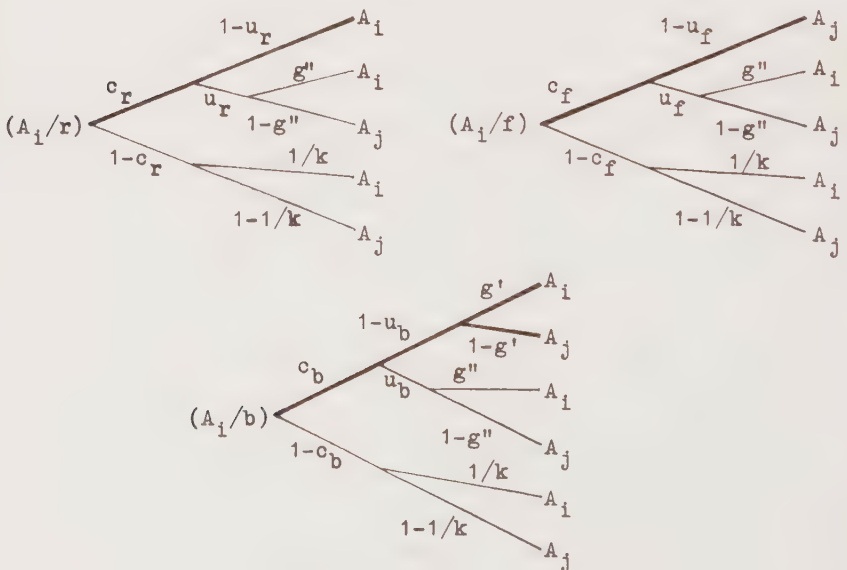


Abb. 4-1: Modifizierte B-Diagramme (s. Text)

Um anzudeuten, daß die Unabhängigkeitspostulate U1 und U2 nicht unbedingt gültig sein müssen, haben c und u Indices (r,f,b) erhalten. In Abschwächung von U3 wird im folgenden von z.B. dem $(c_r, 1-u_r)$ -Weg gesprochen, nicht aber von dem $c_r \cdot (1-u_r)$ -Weg, da ja der Multiplikationspunkt der Ausdruck von U3 ist.

Der Parameter g wurde aufgespalten in g' (Beharrung auf A_i bei Erinnerung an b) und in g'' (Beharrung auf A_i bei Nichterinnerung an $O=r,f;b$; s. S. 30). Falls $g' = g''$ ist (nach A3b (iii) auf S. 29), geht das obige Diagramm (A_i/b) wieder in das auf S. 31 über. Diese Aufspaltung dient hauptsächlich der Vereinheitlichung der Darstellung: Insbesondere soll auch im (A_i/b)-Diagramm die Nichterinnerung an O genauso dargestellt sein wie in den (A_i/r)- und (A_i/f)-Diagrammen.

Prinzipiell unterschiedlich bleibt das (A_i/b)-Diagramm gegenüber den beiden anderen hinsichtlich des $(c, 1-u)$ -Weges. Bei der Erinnerung an b bleibt der Vp die Entscheidung, ob sie ihre Antwort A_i wiederholen (mit g') oder nicht mehr wiederholen will (mit $1-g'$). Diese Freiheit der Entscheidung hat sie bei Erinnerung an r bzw. f nicht.

In allen drei Diagrammen sind die oberen Zweige, also die $(c, 1-u)$ -Wege, "dick" durchgezogen, alle anderen Zweige, also die (c, u) - bzw. $(1-c)$ -Wege, "dünn" gezeichnet. In Analogie zu den "gelernten" bzw. "ungelernten" Zuständen (C_2 bzw. C_1) in dem oben erläuterten "Ein-Element"-Modell (s. S. 62f.) entsprechen die "dicken" Wege dem Zustand, daß R und O erinnert (gelernt) wurden, und die "dünnen" Wege dem Zustand, daß R und O nur teilweise bzw. überhaupt nicht erinnert wurden.

Die (c, u) - bzw. $(1-c)$ -Wege werden also unterschiedslos zusammengefaßt. Eine feinere Betrachtung wäre natürlich möglich, soll aber in dieser Arbeit nicht durchgeführt werden. Diese Vorentscheidung wird so begründet:

Bei Antworten, die über die (c, u) - bzw. $(1-c)$ -Wege zustandekommen, wird die Vp im allgemeinen versuchen, die entfallene(n) Erinnerung(en) an R bzw. O doch noch wiederzufinden. Nach

mehreren Versuchen wird sie schließlich mit großer subjektiver Unsicherheit eine Reaktion abgeben. Ob nun die Antwort richtig oder falsch wird, auf jeden Fall werden im Durchschnitt die Antwortzeiten verhältnismäßig lang sein und stark schwanken. (Zu den Schwankungen nach unten können z.B. auch relativ schnelle "Resignationsantworten" oder subjektiv sichere, aber falsche Antworten beitragen. Dieser Anteil kann aber als klein angenommen werden.)

Die Berechtigung zu differenzierteren Hypothesen bezüglich Unterschieden zwischen (c,u)- und (1-c)-Wegen scheint nicht gegeben zu sein.

Es wird also angenommen, daß alle sechs betrachteten Verteilungen (zu den Medianen $T_{rr} \dots T_{b\bar{b}}$ gehörend) einheitlich eine relativ langsame Grundkomponente enthalten, eben die (c,u)- und (1-c)-Zeiten. Diese Grundkomponente sei mit GK abgekürzt. Die unterschiedlichen Hypothesen ergeben sich nun dadurch, daß dieser GK Zeitanteile zugemischt werden, die über die (c,1-u)-Wege zustande kommen (s. S. 65; s. auch zur Illustration die graphische Darstellung zur Aufspaltung einer rr-Verteilung auf S. 207 in 7.1.3). Über diese Anteile können genauere Annahmen gemacht werden als über GK.

Die Annahmen werden besonders präzise sein, wenn sie sich auf die Erinnerungen und Entscheidungen "längs" der (c,1-u)-Wege beschränken, die mit großer subjektiver Sicherheit verbunden sind. Subjektiv unsichere (c,1-u)-Erinnerungen, obwohl sie korrekt sind, führen wegen des starken Einflusses der subjektiven Sicherheit auf Antwortzeiten (s. 3.4) zu langen Zeiten, die mit denen aus GK vergleichbar sind. (Eine Beleuchtung der quantitativen Verhältnisse kann prinzipiell die Analyse der "Verbalisierungsdaten" und der zugehörigen Zeiten erbringen.)

Es wird daher im weiteren so vorgegangen, als ob eine klare Grenze zwischen der GK und den "sicheren" (c,1-u)-Verteilungen besteht.

In Analogie zum "Ein-Element"-Modell entsprechen also der GK-Median dem zum Zustand C_1 gehörenden Mittelwert μ_1 und die

mittleren "sicheren" $(c, 1-u)$ -Zeiten dem zu C_2 gehörenden Mittelwert μ_2 . Auch die Relation $\mu_2 < \mu_1$ kann analog übertragen werden, bis auf die Modifikationen, die durch Umschaltvorgänge längs der $(c_f, 1-u_f)$ - und $(c_b, 1-u_b, 1-g')$ -Wege entstehen.

In den folgenden Hypothesen sollen die Symbole $T_{rr} \dots$ Populationsmediane bedeuten; sonst stehen sie für die Stichprobenmediane.

Hypothese H1: $T_{rr} = T_{ff}$

Die zu T_{rr} bzw. T_{ff} gehörigen Verteilungen enthalten nur die GK, unterscheiden sich also nicht. Man kann auch sagen: GK wird durch diese beiden Verteilungen definiert. - Sollte sich aus den Daten ein signifikanter Unterschied zwischen T_{rr} und T_{ff} ergeben, müßte GK einer differenzierteren Analyse unterworfen werden.

Hypothese H2: $T_{rr} < T_{rr}$

Die T_{rr} -Verteilung enthält zusätzlich zu GK noch den schnellsten überhaupt vorkommenden Weg $(c_r, 1-u_r)$: Die erinnerte Antwort R kann aufgrund der erinnerten Rückmeldung $O=r$ mit minimalem Entscheidungsaufwand abgegeben werden (s. zur Unterstützung der Hypothese die empirischen Befunde bezüglich T_s und T_e in 3.2, aber auch die einschränkende Anmerkung auf S. 46 hinsichtlich der Vergleichbarkeit von T_{rr} bzw. T_{rr} und T_s bzw. T_e).

Hypothese H2': $T_{rr} < T_{ff}$

Zur Begründung s. H2. Die Hypothese H2' ist besonders von Interesse, falls H1 abgelehnt werden müßte. Sie ist aber auch dann nicht redundant im Verhältnis zu H1 und H2, wenn die Ergebnisse nicht gegen H1 sprechen; denn die als Nullhypothese formulierte Hypothese H1 kann zwar beibehalten, aber nicht bewiesen werden.

Hypothese H3: $T_{rr} < T_{ff}$

Die T_{ff} -Verteilung enthält zusätzlich zu GK den $(c_f, 1-u_f)$ -Weg. Die Erinnerungen an R und $O=f$ können an sich schnell

stattfinden, das Umschalten auf eine andere Antwortalternative erfordert aber mehr Zeit als das einfache Wiederholen von R beim $(c_r, 1-u_r)$ -Weg (s. auch KELLER et al., 1965; ERICKSON et al., 1966; GREENWALD, 1970).

Eine Hypothese bezüglich T_{ff} und z.B. T_{rr} oder T_{ff} kann nicht gut begründet werden, da nicht bekannt ist, in welchem Ausmaß die Schnelligkeit der Erinnerungen an R und O=f durch das anschließende Umschalten eventuell wieder rückgängig gemacht wird. Die Umschaltzeit hängt vermutlich auch von k ab (s. Hypothese H7).

Hypothese H4: $T_{bb} < T_{rr}$

Die T_{bb} -Verteilung enthält zusätzlich zu GK noch den relativ schnellen $(c_b, 1-u_b, g')$ -Weg: Aufgrund einer klaren Instruktion hinsichtlich der Bedeutung von b und im Anschluß an die vorgesehene Einübungsliste kann erwartet werden, daß Vpn in mehr oder weniger starkem Ausmaß eine Vorentscheidung getroffen haben, wie sie sich bei Erinnerung an R und O=b verhalten wollen, und daher nicht längere Zeit für eine Entscheidung brauchen.

Hypothese H4': $T_{bb} < T_{ff}$

Zur Begründung s. H4. Zum Verhältnis von H4' zu H4 und H1 s. die Anmerkung bei H2'

Hypothese H5: $T_{rr} < T_{bb}$

Der $(c_b, 1-u_b, g')$ -Anteil innerhalb der T_{bb} -Verteilung enthält eine zusätzliche, wenn auch u.U. nur kurze Entscheidungszeit, die aber als länger anzunehmen ist als die Minimalentscheidungszeit beim $(c_r, 1-u_r)$ -Weg (s. H2 und H4).

Es ist vorstellbar, daß einige Vpn eine so strikte Vorentscheidung bezüglich ihres Verhaltens nach b getroffen haben, daß ihre Antwortzeiten längs der $(c_r, 1-u_r)$ - und der $(c_b, 1-u_b, g')$ -Wege vergleichbar sind. Es ist aber unwahrscheinlich, daß dieses für alle Vpn zutrifft.

Hypothese H6: $T_{bb} < T_{b\bar{b}}$

Die Erinnerungen an R und O=b können längs des $(c_b, 1-u_b)$ -Weges für beide Verteilungen gleich schnell erfolgen, das Umschalten $(1-g')$ auf eine andere Antwortalternative erfordert aber mehr Zeit als das einfache Wiederholen (g') . - Eine Hypothese bezüglich $T_{b\bar{b}}$ und $T_{r\bar{r}}$ oder T_{ff} kann, wie auch bei T_{ff} , nicht begründet werden, da das Ausmaß des Umschalteffekts unbekannt ist. Die Umschaltzeit hängt vermutlich auch von k ab (s. Hypothese H8).

Die bisherigen Hypothesen bezogen sich sowohl auf k=2- als auch k>2-Aufgaben. Mit den folgenden drei Hypothesen soll eine etwaige k-Abhängigkeit der Mediane erfaßt werden.

Hypothese H7: $(T_{ff} - T_{rr})_{k=2} < (T_{ff} - T_{rr})_{k>2}$

Es wird Bezug genommen auf die zusätzliche Umschaltzeit bei Erinnerung an O=f gegenüber der Minimalentscheidungszeit bei Erinnerung an O=r (s. H3; für die Aufstellung von H7 wird also die Bestätigung von H3 vorausgesetzt). Während im Fall k=2 bei Erinnerung an R und O=f die Entscheidung, die andere Antwortalternative zu nennen, leicht ist und sehr schnell erfolgen kann, ist im Fall k>2 die Entscheidung, eine wenn auch beliebige von den übrigen k-1 Antwortmöglichkeiten zu sagen, schwieriger und dauert demnach länger (s. zur Unterstützung der Hypothese die allgemeine Abhängigkeit der Latenzzeiten von der Antwortalternativenanzahl; S. 56f.). H7 stellt gewissermaßen eine Präzisierung von H3 dar.

Hypothese H8: $(T_{b\bar{b}} - T_{bb})_{k=2} < (T_{b\bar{b}} - T_{bb})_{k>2}$

Für die Aufstellung von H8 wird die Bestätigung von H6 vorausgesetzt. Zur Begründung s. H7. Durch H8 wird H6 präzisiert.

Hypothese H9: $T_{k=2} < T_{k>2}$

Mit $T_{k=2}$ bzw. $T_{k>2}$ ist der jeweilige Gesamtmedian aller sechs Mediane bezeichnet. Während in H7 und H8 der Umschalteffekt

in Abhängigkeit von k isoliert, nämlich durch die Differenzbildung, betrachtet wurde, bezieht sich H_9 global auf die allgemeine k -Abhängigkeit von Antwortzeiten. Die Hypothese H_9 soll nicht in sechs Einzelhypothesen bezüglich $T_{rr}...$ T_{bb} aufgespalten werden, also nicht über ihre theoretische Bedeutung hinaus aufgebläht werden. Falls sich Signifikanz ergeben sollte, wird der Ursprung wohl hauptsächlich bei T_{fr} und auch T_{bb} liegen (s. H_7 und H_8). Es sei aber auch an die Befunde von HALL (1969) erinnert: $T_{rr}(k=2) < T_{rr}(k>2)$ und $T_{rr}(k=2) < T_{rr}(k>2)$.

Zu beachten ist, daß die obigen Hypothesen im wesentlichen unabhängig voneinander begründet wurden, insofern sie sich auf verschiedene Wege in den B-Diagrammen beziehen. H_2 und H_2' (bzw. H_4 und H_4') lassen sich z.B. durchaus unabhängig voneinander aufstellen, ungeachtet des Hinweises auf H_1 . Die Hypothese H_9 ist allerdings schwach abhängig von H_7 und H_8 . Da etwaige Tests der einzelnen Hypothesen an teilweise denselben Daten durchgeführt werden, sind sie schon allein aus diesem Grunde voneinander abhängig. Die Auswirkungen auf das Signifikanzniveau werden im Auswertungsteil von 5.2.2 diskutiert werden.

In später zu berichtenden Experimenten werden mehrere Lernlisten dargeboten werden (in den Positionen I und II), d.h. es wird mit Übungs- bzw. Ermüdungseffekten zu rechnen sein. Inwieweit sich dieser Positionsfaktor auf die mittleren Latenzzeiten auswirkt, ist aufgrund der Literatur nicht annähernd vorausszusagen. Es ist ebenso plausibel anzunehmen, daß durch die Gewöhnung der Vpn an die unbekannte Experimentalsituation die Antwortzeiten kürzer werden, als auch, daß die Vpn den Anfangsstress allmählich überwinden und daraufhin ruhiger antworten (s. dazu auch BERGHÄUSER u. SCHRÖDER, 1972, S. 24).

Falls die Zeitdaten keinen auffälligen Positionseffekt zeigen, sollen die Lernlisten zusammengelegt und gemeinsam analysiert werden.

Über H1 bis H9 hinaus konnten keine weiteren Hypothesen gut begründet abgeleitet werden.

Inwieweit werden nun schon durch diesen Hypothesenkomplex die theoretisch möglichen Rangreihenfolgen der sechs Mediane ("Latenzzeitmuster") eingeschränkt? - Es sollen der Einfachheit halber nur H1 bis H6 besprochen werden, da sie keinen Bezug auf die k-Abhängigkeit nehmen. - Sollte tatsächlich überhaupt keine Abhängigkeit der Mediane von den Rückmeldungsmodalitäten existieren, so gäbe es $6! = 720$ gleichwahrscheinliche Reihenfolgen der sechs Mediane (numerische Gleichheit zwischen zwei experimentell gefundenen Medianen kann ausgeschlossen werden, da auf der kontinuierlichen Zeitskala gemessen wird). Aber nur 30 von diesen 720 Zeitmustern sind auch noch mit dem Hypothesenkomplex H1 bis H6 verträglich, wie man sich durch systematisches Hinschreiben der zulässigen Reihenfolgen überzeugen kann.

Der optimale Fall wäre der einer einzigen zulässigen Reihenfolge, wie sie durch ein mathematisches Latenzzeitmodell vorhergesagt werden könnte.

Zusammenfassend ist zu sagen, daß die eventuell positive Beantwortung einzelner Hypothesen aus H1 bis H9 relativ wenig besagt. Erst in ihrer Gesamtheit beleuchten sie, gemäß Fragestellung (2), die (In-)Adäquatheit der Schlußfolgerungen, die aus dem BUCHWALDSchen Ansatz bezüglich der Latenzzeiten gezogen wurden. Es soll nochmals betont werden, daß diese Schlußfolgerungen in möglichst einheitlicher Weise aus den B-Diagrammen abgeleitet wurden.

5. Voruntersuchung: Test des BUCHWALD-Modells

Im Kapitel 4 bezogen die Hypothesen H1 bis H9 sich auf die allgemeinen Fragestellungen (1) und (2), wobei die strikte Gültigkeit des BM nicht vorausgesetzt wurde und H1 bis H9 somit sowohl auf die jetzt zu behandelnden Voruntersuchungen als auch auf die späteren Hauptuntersuchungen anwendbar sind. In diesem Kapitel soll sich im wesentlichen auf die Fragestellung (2'a) beschränkt werden. Es wird im Experiment I das BM (1969) daraufhin getestet, ob es als hinreichend sichere Ausgangsbasis für ein darauf aufbauendes Latenzzeitmodell dienen kann. Vorher sollen in 5.1 einige Modifikationen der BUCHWALD-schen Versuchsanordnung besprochen werden.

Im Experiment II wird diese modifizierte Versuchsanordnung übernommen; die Konzentration liegt besonders auf dem Einfluß der Instruktion.

5.1 Modifikationen der BUCHWALD-schen Versuchsanordnung

5.1.1 Beobachtungslernen

Wie in 3.2 erwähnt und versucht zu erklären, hatten vergleichbare Arbeiten (BERGHÄUSER u. SCHRÖDER, 1972; HEIDE, 1974) keine interpretierbaren Unterschiede zwischen den Zeitmedienen erbracht. Eine weitere mögliche Erklärung wäre, daß das Lernniveau der Vpn zu gering war. Folglich wären nur sehr geringe Anteile der Zeitverteilungen über die (c,1-u)-Wege zustandegekommen. Die beobachteten Verteilungen bzw. die Mediane drückten dann hauptsächlich die allen Bedingungen rr, ... b5 gemeinsame GK aus.

Das Problem war also, ein höheres Lernniveau zu erzeugen, ohne die BUCHWALD-sche Versuchsanordnung in ihren wesentlichen Zügen zu sehr zu verändern. Ein höheres Lernniveau erhält man üblicherweise dadurch, daß man die Vpn mehrere Lern-durchgänge absolvieren läßt (s. z.B. THORNDIKE). Aber gerade

dieses Vorgehen hatte THORNDIKE die Kritik eingebracht, daß in den ersten Durchgängen die Rückmeldung f, in den späteren Durchgängen r überwog, also Akzentuierungs-(Isolations-)Effekte auftreten konnten (s. S. 14). Aus diesem Grunde waren in den BUCHWALDschen Lernlisten die Rückmeldungen gleich häufig festgelegt, und es fand auch nur ein Lerndurchgang statt. Das Problem, mehrere Lerndurchgänge zu gestatten, ohne dabei die Rückmeldungsproportionen zu verändern, wurde nun folgendermaßen gelöst:

Die Lernliste wurde nicht mehr Wort für Wort dargeboten, wobei die Vp zu jedem Reizwort, z.B. SIEB, eine Reaktion, z.B. 3, abgab und daraufhin eine Rückmeldung, z.B. richtig, vom V1 erhielt, sondern das Lernmaterial wurde der Vp auf Kärtchen (weiß, DIN A 7) geschrieben vorgelegt. Auf einem Kärtchen stand z.B.

S I E B 3 richtig,

auf einem anderen z.B.

T E I G 7 falsch,

wiederum auf einem anderen

W E I N 5 — , usw.

Dabei bedeutete der Bindestrich "—", daß der Vp keine Information gegeben wurde, ob nun z.B. 5 eine richtige oder falsche Antwort auf WEIN wäre; "—" ist also gleichbedeutend mit der Rückmeldung b.

Entsprechend der Anzahl (= 40) der Reizworte in der Lernliste gab es 40 Lernkärtchen (12 r-, 12 f- und 12 b-Kärtchen; zuzüglich je zwei "Puffer"-Kärtchen am Anfang und Ende der Liste). Das weitere Vorgehen kann aus der nun folgenden, sehr ausführlichen Hauptinstruktion entnommen werden. Auf einige Besonderheiten wird anschließend eingegangen.

Hauptinstruktion zur "Standard"- bzw. "R-corr"-Bedingung (s. S. 32 u. 86) in Experiment I (beispielsweise für die Versuchsbedingung $k = 4$):

"Du nimmst jetzt an einem I e r n versuch teil, dessen Ergebnisse von Bedeutung sind für den Programmierten Unterricht.

Zum besseren Verständnis des Hauptversuchs, an dem Du teilnimmst, erkläre ich Dir im folgenden einen Vorversuch, an dem eine andere Versuchsperson vor Dir teilgenommen hatte.

In diesem Vorversuch wurden einer Versuchsperson (Vp) eine Reihe einsilbiger deutscher Wörter gezeigt, denen vorher vom Versuchsleiter (Vl) die Zahlen 3, 5, 7 oder 9 z u f ä l l i g und gleich häufig zugeordnet waren. Die Vp sollte nun lernen, welche Zahl zu welchem Wort gehört: sie sah ein Reizwort und antwortete mit einer der Zahlalternativen, wobei sie zunächst nur durch Raten die richtige Zuordnung treffen konnte. Anschließend erfuhr sie sofort, ob die von ihr getroffene Zuordnung von Zahl zu Reizwort r i c h t i g oder f a l s c h war, oder es erfolgte k e i n e Stellungnahme von seiten des Vl darüber, ob ihre Antwort richtig oder falsch war. Beispiele für den L e r n d u r c h g a n g :

	Reizwort	Antwort der Vp	Rückmeldung des Versuchsleiters
Lerndurchgang	S I E B	3	richtig
	T E I G	7	falsch
	W E I N	5	—

(wobei "—" bedeutet: keine Stellungnahme, ob die Antwort richtig oder falsch ist. Aus der englischen Fachliteratur über ähnliche Versuche wurde hierfür der Terminus blank übernommen.)

Die Aufgabe der Vp war es, im P r ü f d u r c h g a n g , in dem jeweils ein Reizwort dargeboten wurde, möglichst wenig Fehler zu machen! Beispiel:

	Reizwort	Antwort der Vp	Rückmeldung des Vl
Lerndurchgang	T E I G	7	falsch
Prüfdurchgang	T E I G	7	

Die Vp hat ihren Fehler aus dem Lerndurchgang wiederholt, sie hat also nicht gelernt!

H A U P T V E R S U C H

Jede Folge von Reizwort-Antwort-Rückmeldung aus dem Lerndurchgang der vorherigen Vp wurde für den jetzigen Hauptversuch auf jeweils eine Karte protokolliert (s. Beispielkarten). Es soll nun untersucht werden, welche Methode sich günstiger auf die Lernleistung auswirkt:

Methode 1, die im Vorversuch gewählt wurde:

Die Vp wählte die Antwort (Zahlalternative)
s e l b s t aus und erfuhr eine entsprechende
Rückmeldung des Versuchsleiters.

Methode 2, die im jetzigen Hauptversuch angewandt wird:

Du l i e s t die protokollierten Folgen von
Reizwort-Antwort-Rückmeldung der Vp vor Dir.

Der Prüfdurchgang ist für beide Methoden gleich: Es wird jeweils ein Reizwort gezeigt. Die Vp soll bei ihren Antworten möglichst wenig Fehler machen!

Der Hauptversuch läuft für Dich konkret folgendermaßen ab:

Du erhältst vom Versuchsleiter eine Reihe von Karten, auf denen je eine Folge von Reizwort-Antwort-Rückmeldung zu sehen ist. Du sollst in dem vom Metronom vorgegebenen Rhythmus eine Folge laut lesen (beachte "—": sprich: blank), dann die Karte beiseite legen und die nächste Folge auf der nächsten Karte lesen, usw., bis zum Ende der Liste.

(Das rhythmische Lesen wird vorher mit neutralen Karten geübt. Auf ihnen erkennst Du drei verschiedene Farbpunkte, die Du im Rhythmus des Metronoms laut benennen sollst, so lange, bis Du das Tempo und das Hantieren mit den Karten beherrscht.)

Du wirst die gleiche Liste in drei Lerndurchgängen hintereinander kennenlernen. Bitte lasse Dich durch das vorgegebene Tempo und die veränderte Reihenfolge der Karten im zweiten Lerndurchgang nicht verwirren oder entmutigen: Du lernst hierbei mehr als Du vermutest!

Solltest Du jedoch einmal aus dem Takt kommen, lasse einen

oder zwei Schläge des Metronoms aus und fahre dann zügig fort.

In dem sich anschließenden P r ü f durchgang wird kein Rhythmus eingehalten. Der Versuchsleiter zeigt Dir durch die Öffnung in der Mitte des vor Dir stehenden Schirms jeweils ein Reizwort.

Deine Aufgabe wird es sein, bei Deinen Antworten (Zahlalternativen) möglichst wenig Fehler zu machen! Du mußt auf jeden Fall eine Antwort geben, eventuell indem Du rätst! Eine Rückmeldung wird nicht mehr gegeben!

Um dem Versuchsleiter das Protokollieren zu erleichtern, werden Deine Antworten auf Tonband aufgenommen. Bitte vermeide deshalb während des Prüfdurchgangs alle weiteren Äußerungen.

Der hier beschriebene Versuchsablauf wird mit mehreren, verschiedenen langen Listen durchgeführt. Um eine Verwechslung zwischen ihnen zu vermeiden, wirst Du Dich zur Entspannung mit einem kleinen Spiel beschäftigen.

Hast Du noch Fragen?

Wiederhole bitte in eigenen Worten, wie der Versuch für Dich ablaufen wird!"

(Ende der Instruktion)

Selbstverständlich handelt es sich bei dem angebotenen Lernmaterial um ein fiktives Lernprotokoll "der Vp vor Dir" (s. dazu S. 82, Punkt 5).

Für den augenblicklichen Stand der Diskussion ist zunächst wichtig, daß also drei Lerndurchgänge gestattet waren, bei denen sich aber die Rückmeldungsproportionen ($r:f:b=12:12:12$) nicht änderten.

Inwieweit ist nun das experimentelle Vorgehen, wie durch die Instruktion beschrieben, speziell mit dem BUCHWALDschen Ansatz und allgemein mit der Frage nach der relativen Effizienz von positiver und negativer Rückmeldung verträglich? Dieses relativ "passive" Lernen durch "Beobachtung" des in diesem Falle protokollierten Lernverhaltens eines anderen muß noch

näher besprochen werden. Dabei braucht allerdings nicht intensiver auf die grundlegenden Arbeiten von BANDURA (1962), BANDURA u. WALTERS (1963) zum Nachahmungs-(Beobachtungs-) Lernen eingegangen zu werden.

Es genügt z.B. die Darstellung in HILGARD u. BOWER (1966, S. 641-645); dort heißt es abschließend: "Außerdem unterscheidet sich BANDURAs experimentelle Situation - wie der Autor selbst bemerkt - eigentlich nicht allzusehr von dem typischen Laboratoriumsexperiment über verbales Lernen, bei dem die Versuchsperson auf einer 'Gedächtnistrommel' (memory drum) dargebotene sinnlose Silben beobachtet und später rezitiert." Es werden daher im weiteren nur Arbeiten zitiert, die Beobachtungslernen in einfachen PAL- oder CI(concept identification)-Experimenten durchführen.

In einer Gruppe von Untersuchungen (z.B. SPENCE, 1970; LIU, 1972) wird Beobachtungslernen als Hilfsmittel eingeführt, um bestimmte experimentelle Bedingungen erzeugen zu können, wie z.B. in vorliegender Arbeit die Erhaltung der Rückmeldungsproportionen. Über diese Hilfsfunktion hinaus wird es nicht weiter diskutiert.

Bei einer anderen Gruppe von Autoren spielt die Gegenüberstellung von "aktivem" und "beobachtendem" Lernen ("performing" vs. "observing") eine bedeutsame theoretische Rolle (z.B. HILLIX u. MARX, 1960; MARX u. MARX, 1970; WITTER et al., 1971; MARX u. WITTER, 1972). Es geht dabei u.a. um die Frage, unter welcher der beiden Lernbedingungen besser gelernt wird, und um das Problem der Trennung von Information und Motivation beim Verstärkungsvorgang.

Auf diese Ansätze wird unter 6.3.3 erneut einzugehen sein; für das Experiment III hat das Beobachtungslernen dann einen höheren Stellenwert als für die augenblicklich behandelten Experimente I und II.

Vorerst, d.h. für die Prüfung des BM, genügt es festzustellen, daß aus diesen Arbeiten kein Einwand abzulesen ist, der die Einführung von Beobachtungslernen suspekt machen könnte. Vielmehr sprechen WITTER et al. (1971), unter einschränkenden

Bedingungen zwar, von einer Vergleichbarkeit der Prozesse bei "aktivem" und "beobachtendem" Lernen.

Bei näherer Betrachtung scheint das Beobachtungslernen nach der "Kärtchen"-Methode sogar dem Zweck der aktuellen Untersuchung dienlicher zu sein als das konventionelle aktive Lernen:

1. Es wird die Gleichhäufigkeit der Rückmeldungen über mehrere Lerndurchgänge hinweg garantiert; das war ja das Ausgangsproblem.

2. Es kann auch die Gleichhäufigkeit der Antworten, z.B. 3, 5, 7 und 9, erzwungen werden (s. S. 29: A3 a). Dadurch wird der Faktor $1/k$ in den Gleichungen (1a) bis (1c) annehmbarer (s. auch LIU, 1972).

3. Das "beobachtende" Lernen ist sicherlich mit weniger Emotionen aufgrund der Rückmeldungen verbunden als das "aktive" Lernen. Daher ist es aber auch besonders für die Untersuchung eines Modells geeignet, das, wie das BM, den unemotionalen, informativen Aspekt der Rückmeldungen hervorhebt.

4. Da die gesamte Sequenz S-R-O von der Vp allein gelesen bzw. gesprochen wird, können die, als per Kontiguität entstehend angenommenen, S-R- und S-O-Verbindungen im Sinne des BM als gleichartiger angenommen werden als z.B. beim Aussprechen von R durch die Vp und der vokalen Rückmeldung O durch den Vl.

5. Man kann a priori die Zusammenstellung der Reizworte, Zahlalternativen und Rückmeldungen so arrangieren, daß keine starken vorexperimentellen "Eselsbrücken" ausgelöst werden. Es besteht ferner die Möglichkeit, geeichtes Lernmaterial herzustellen, geeicht u.a. hinsichtlich des "Assoziabilitätswertes" (associability value AS; s. z.B. MONTAGUE u. KIESS, 1968). Diese letztere Möglichkeit wurde aber hier aus Zeitgründen nicht weiter verfolgt.

5.1.2 Weitere Modifikationen

Aus der Instruktion auf S. 77-80 gehen weitere Modifikationen der BUCHWALDschen Versuchsanordnung hervor.

Eine der wichtigsten Voraussetzungen für die strikte Gültigkeit des BM war die Verhinderung von Rehearsal. Diese wurde im Experiment I dadurch angestrebt, daß die Vp im relativ schnellen Rhythmus eines Metronoms die Kärtchen lesen und wechseln sollte; pro Kärtchen standen 2,6 sec zur Verfügung. Wenn diese Manipulation erfolgreich war, sollte die Wirkung der drei Lerndurchgänge nur darin bestehen, entsprechend U1 und U2 (S. 34), $c_r = c_f = c_b$ und $1 - u_r = 1 - u_f = 1 - u_b$ gleichmäßig anzuheben, nicht aber z.B. c_r stärker anwachsen zu lassen als c_f und c_b .

Das rhythmische Lesen und auch Symmetriegründe bezüglich der mathematischen Gleichbehandlung der drei Rückmeldungen r, f und b erforderten eine weitere Modifikation. War nach der vorgeplanten Ordnung in einer Liste eine b-Rückmeldung an der Reihe, also "Schweigen" seitens des V1 bei BUCHWALD, mußte die Vp explizit "blank" laut aussprechen, d.h. also: blänk!

In Anbetracht der Anmerkungen zur experimentellen Darstellung von b (S. 21) und der klaren Instruktion bezüglich der Bedeutung von "blank" schien diese Abänderung erlaubt zu sein.- Auch sind die Ergebnisse aus den folgenden Experimenten I bis V so, daß kein Anlaß besteht, an der Vergleichbarkeit der beiden Darstellungen von b zu zweifeln. Daß die b-Rückmeldung, ob nun "Schweigen" oder "blank", an sich in Frage gestellt werden soll (s. 6.2.3), ist ein anderes Problem.

Abgesehen davon, daß die hier gebrauchte Instruktion länger und komplizierter ist (wegen der "Kärtchen"-Methode) als die bei BUCHWALD verwendete, besteht eine sehr feine unterscheidende Nuancierung hinsichtlich der Betonung der Rückmeldung f. Bei BUCHWALD heißt es allgemein (s. S. 22): "that his task was to learn which number went with which word."

In der vorliegenden Instruktion wird als Beispiel für den

Prüfdurchgang ein f-Item gebracht, und es wird gesagt, daß eine Wiederholung nach f bedeutet, daß nicht gelernt worden ist.

Dieser feine Unterschied ist im Zusammenhang mit dem Experiment I wohl vernachlässigbar, er wird aber später im Experiment II mehr betont und wesentlich werden.

Eine weitere leichte Modifikation bezieht sich auf die Durchführung der "reminder condition" und der "delayed outcome"-Bedingung (S. 34 u. 36). Nach BUCHWALD behandeln beide Bedingungen das Problem, im Prüfdurchgang sichere Erinnerungen an R bzw. O zu erzeugen, und zwar in "analoger" Weise. Das letztere ist nicht ganz richtig, denn unter der "reminder condition" wird im Prüfdurchgang an die Antwort R erinnert, die schon im Lerndurchgang einmal ausgesprochen worden war, während unter der "delayed outcome"-Bedingung im Prüfdurchgang die Rückmeldung O angegeben wird, die der VL im Lerndurchgang hätte sagen können, aber nicht gesagt hat. Insofern handelt es sich nicht um eine zur "response reminder condition" analoge "outcome reminder condition".

Um beide Bedingungen einheitlicher behandeln zu können (auch hinsichtlich der speziellen Instruktionen), wurde die "delayed outcome"-Bedingung in eine "O-remind"-Bedingung abgewandelt, d.h. es wurden also auch schon im Lerndurchgang die Rückmeldungen O angegeben, an die im Prüfdurchgang erinnert wurde. Für die Gleichungen (3a) bis (3c) auf S. 36 hat diese experimentelle Manipulation keine Folgen, wie es auch im Rahmen des BUCHWALDSchen Informationsansatzes sein sollte. Die "reminder condition" sei dementsprechend jetzt in "R-remind"-Bedingung umbenannt; s. S. 35.

Der Charakter einer "verzögerten Rückmeldung", wie er z.B. bei der Kontrastierung zu Verstärkungstheorien wichtig ist, geht durch obiges Vorgehen allerdings zu einem gewissen Teil verloren. Da aber diese Diskussion schon von BUCHWALD, eben durch die "delayed outcome"-Bedingung, qualitativ und für sein Modell erfolgreich behandelt wurde, sollen im Experiment I die beiden vereinheitlichten "remind"-Bedingungen, also

R-remind und O-remind, durchgeführt werden.

Neben vielen kleineren Abänderungen der BUCHWALDschen Versuchsanordnung, die schon allein durch die Darbietungstechnik ("Kärtchen"-Methode) bedingt sind, ist noch auf folgende zwei explizit hinzuweisen:

1. Es wurden Latenzzeiten gemessen; das entspricht ja dem Ziel dieser Arbeit.
2. Anschließend an den letzten Prüfdurchgang wurde ein festgelegtes Interview mit den Vpn durchgeführt. In Anbetracht dessen, daß BUCHWALD bei Aufstellung seines Modells u.a. die postexperimentellen Aussagen von Vpn zugrundelegte (S. 27f.), schien es nützlich, möglichst modellrelevante Fragen in den Interviewkatalog aufzunehmen (s. S.110f.).

5.2 Experiment I: Zum BUCHWALD-Modell

5.2.1 Fragestellung und Hypothesen

Das Ziel dieses Experiments ist es, gemäß Fragestellung (2'a), das BM daraufhin zu testen, ob es als hinreichend bewährte Ausgangsbasis für ein darauf aufzubauendes Latenzzeitmodell dienen kann. Im wesentlichen wird dabei dem Vorgehen BUCHWALDs (1969) gefolgt, d.h. es werden die dort angegebenen Tests ("reminder condition"; "delayed outcome", "apparent recall") durchgeführt, teilweise in abgeänderter oder erweiterter Form.

Ferner werden die unter 5.1 eingeführten Modifikationen, insbesondere Beobachtungslernen und explizit ausgesprochenes "blank", erprobt. Zu den Abänderungen gehört auch die Messung der Latenzzeiten, die es ermöglicht, die in Kapitel 4 aufgestellten Hypothesen H1 bis H9 zu überprüfen.

Im Hinblick auf die in 2.3.2 postulierten unterschiedlichen Strategien in Abhängigkeit von der Antwortalternativenzahl k werden in diesem Experiment, wie auch in den meisten anderen

der hier zu referierenden Versuchsgruppen, $k=2$ und als Vertreter für $k>2$ -Bedingungen $k=4$ untersucht (s. BUCHWALD, 1962).

Im folgenden werden die Hypothesen durch ein "h" gekennzeichnet, im Unterschied zu den Latenzzeithypothesen H1 bis H9. Die einzelnen Hypothesen, h1 bis h15, sind nicht alle gleich relevant, tragen aber alle zur Beurteilung der Adäquatheit des BM bei (s. dazu auf S. 96: Zur statistischen Auswertung).

R-corr: Hypothesen zu den Gleichungen (1a) bis (1c) auf S. 32

Die zu diesen Hauptgleichungen des BM gehörende Versuchsdurchführung sei im folgenden Standardbedingung oder "R-corr"-Bedingung genannt. Die letztere Bezeichnung dient der einheitlichen Kennzeichnung der unterschiedlichen Versuchsgruppen (s.u.); "R-corr" bezieht sich darauf, daß die Reaktion R aus dem Lerndurchgang im Prüfdurchgang unter Umständen - nämlich nach f - korrigiert werden muß. Trotz dieser Bezeichnung, die das Nichtwiederholen einer Antwort hervorhebt, ist die R-corr-Bedingung natürlich unter "Wiederholungsprüfung" zu klassifizieren (s. S. 38).

h1: Unter der R-corr-Bedingung ist zunächst folgende Hypothese zu prüfen: Die mit Hilfe der "Kärtchen"-Methode (Beobachtungslernen!) erhobenen Daten unterscheiden sich nicht wesentlich von den unter "aktiven" Lernbedingungen erhaltenen Daten (z.B. bei BUCHWALD, 1962).

Für die Bestätigung von h1 genügt es, wenn die Profile, die sich aus den unter der "Kärtchen"-Methode gefundenen Ergebnissen errechnen lassen, denen auf S. 24 einigermaßen ähnlich sind. D.h., sie sollten monoton von rr über bb nach ff abfallen und auch die typische Abhängigkeit von k zeigen. Im übrigen ist zu erwarten, daß durch die drei Lerndurchgänge ein höheres Lernniveau, d.h. steilere Profile, erzeugt wird.

h2: Da im Verlauf des Experiments zwei verschiedene Lernlisten, A und B, nacheinander zu lernen sind, gibt es unter der R-corr-Bedingung zwei Übungsstadien (Positionen I und II). Diese seien durch R-corr(I) und R-corr(II) bezeichnet. Zu-

sätzlich zu h_1 ist für R-corr (II) noch folgende Hypothese aufzustellen: Die Profile unter R-corr(II) fallen steiler ab als die unter R-corr(I). - Es wird also vorausgesetzt, daß der Übungseffekt stärker als der Ermüdungseffekt sein wird.

h_3 : BUCHWALD (1969) selbst hat keinen unmittelbaren Test der Gleichungen (1a) bis (1c) vorgenommen. Das ist verständlich, denn es sind drei Parameter zu schätzen (c,u,g), und es sind auch nur drei Gleichungen für diesen Zweck vorhanden. Man kann aber überprüfen, ob die Parameter gegenüber experimentellen Manipulationen, z.B. Veränderung von k, einigermaßen invariant sind. Eine Hypothese dazu: Die aus $k=2$ -Daten geschätzten Parameter unterscheiden sich nicht wesentlich von den aus $k=4$ -Daten geschätzten Parametern. - Andernfalls verliere das BM die elegante Einfachheit seiner Hauptgleichungen.

R-remind: Hypothesen zu den Gleichungen (2a) bis (2c) auf
S. 35

Wie in 5.1.2 nahegelegt, soll im weiteren statt von "reminder condition" von der "R-remind"-Bedingung gesprochen werden, weil eben im Prüfdurchgang an die Antwort R vom V1 erinnert wird.

Die R-remind-Bedingung wird "gemischt" durchgeführt, d.h. in den Lernlisten befinden sich sowohl R-corr-Items als auch R-remind-Items. (Die Vpn erfahren den Typ der Items natürlich erst im Prüfdurchgang.) Aus den ersteren lassen sich die Differenzen (1a)-(1b) schätzen, aus den letzteren die Differenzen (2a)-(2b).

h_4 : Es lautet dann eine Hypothese (s. S. 35): Die Differenz (2a)-(2b) ist größer als die Differenz (1a)-(1b).

h_5 : Eine weitere Hypothese und der zugehörige Test beziehen sich auf die Vorhersage von (2a) aus (2b) und (2c) (s. S. 35).

h_6 : Wiederum kann auch die Invarianz der Parameter (u,g) in Abhängigkeit von $k=2$ bzw. $k=4$ überprüft werden.

O-remind: Hypothesen zu den Gleichungen (3a) bis (3c) auf
S. 36

Wie in 5.1.2 erläutert, wurde die "delayed outcome"-Bedingung nach BUCHWALD in eine "O-remind"-Bedingung abgewandelt. Die O-remind-Bedingung wird ebenfalls "gemischt" (d.h. R-corr-Items u n d O-remind-Items in einer Liste) durchgeführt.

h7: Nach S. 36 ergibt sich als Hypothese: Die Differenz (3a)-(3b) ist größer als die Differenz (1a)-(1b).

Von BUCHWALD selbst nicht betrachtete Testmöglichkeiten sind:

h8: (1c) und (3c) sind bis auf Zufallsschwankungen gleich.

h9: Aus (3a) und (3b) ergeben sich, entsprechend U1 (S. 34), angenähert gleiche Parameterschätzungen für c.

h10: Ferner sollte die Schätzung von c invariant gegenüber der Bestimmung aus k=2- oder k=4-Daten sein.

R-recall: Hypothesen zu der Gleichung (4) auf S. 37

Im weiteren soll statt von der "apparent recall"- von der "R-recall"-Bedingung gesprochen werden, weil die Vp sich im Prüfdurchgang an ihre Antwort R aus dem Lerndurchgang erinnern und sie sagen soll. Um unter dieser Bedingung die Erinnerungen an R möglichst sauber isoliert untersuchen zu können, wurde abweichend von BUCHWALD nur ein "R-recall"-Prüfdurchgang nach den Lerndurchgängen gegeben. Es wurden also keine Standardprüfdurchgänge durchgeführt. Dadurch entfällt die Möglichkeit, die Wahrscheinlichkeiten nach (4) mit denen nach (1a) bis (1c) zu vergleichen, was aber wohl kein großer Verlust ist.

h11: Wichtiger ist, die Unabhängigkeitsannahme U1: $c_r = c_f = c_b = c$ testen zu können. D.h. also, daß Antworten, auf die im Lerndurchgang r, f oder b folgte, unabhängig von diesen Rückmeldungen gleich gut erinnert werden sollten.

h12: Wiederum kann auch die Invarianz von c gegenüber Änderung von k getestet werden.

O-recall: Hypothesen zu der Gleichung (5) (s. unten)

Analog zur R-recall-Bedingung kann auch eine "O-recall"-Bedingung durchgeführt werden. Nach den Lerndurchgängen wird die Vp aufgefordert, sich an die Rückmeldungen O aus den Lerndurchgängen zu erinnern und sie zu sagen. (Wie auf S. 38 erwähnt, ist BUCHWALD nicht weiter auf diese spezielle Bedingung eingegangen.) Als Wahrscheinlichkeit für die Erinnerung an O ergibt sich aus dem BM, analog zur Gleichung (4):

$$P(\text{"O-recall"}) = (1-u) + u/3 \quad (5)$$

(Der Faktor $1/3$ ergibt sich aus der Tatsache, daß r, f und b in den Lerndurchgängen gleich häufig auftreten. Für O spielt also $1/3$ dieselbe Rolle wie $1/k$ für R.)

h13: Gemäß der Unabhängigkeitsannahme U2: $1-u_r=1-u_f=1-u_b=1-u$ gilt (5) sowohl für r-, f- und b-Items getrennt als auch für alle zusammen. Es ist also zu testen, ob die Rückmeldungen r, f und b gleich gut erinnert werden.

h14: Ferner sollte der Parameter u (bzw. $1-u$) invariant gegenüber k sein.

Die R- bzw. O-remind- und die R- bzw. O-recall-Bedingungen werden nur in Position II durchgeführt werden.

Vergleicht man die hier aufgestellten Hypothesen und die zugehörigen Tests mit denen von BUCHWALD (1969), so ist festzustellen, daß BUCHWALD eine Reihe von Testmöglichkeiten bezüglich seines Modells außer acht gelassen hat.

Eine weitere Besonderheit des Experiments I gegenüber den in BUCHWALD (1969) referierten Versuchen ist, daß stärker versucht wurde, die allgemeinen Versuchsbedingungen für die verschiedenen Gruppen (R-corr(II), R-remind, O-remind, R-recall, O-recall) bis auf die beabsichtigten Manipulationen möglichst konstant zu halten (s. S. 91: Versuchsplan).

h15: Dadurch wird der Parameterinvarianz z w i s c h e n den Testgruppen eine größere Chance gegeben. Es sollten z.B.

Schätzungen von u aus R-remind-Daten (s. Gleichungen (2b) und (2c) auf S. 35) und aus O-recall-Daten (s. Gleichung (5) auf S. 89) vergleichbar sein.

Die Forderung der Parameterinvarianz stellt ein mathematisches Modell auf eine sehr harte Probe (s. z.B. ATKINSON et al., 1965, S. 19).

Die für die Latenzzeiten in Kapitel 4 aufgestellten Hypothesen H1 bis H9 gelten unabhängig vom Ausfall der obigen Hypothesentests. Sie werden ausschließlich unter der R-corr(I)-Bedingung behandelt werden, da nur diese Gruppe genügend viele Vpn enthält.

BUCHWALD (1966) berichtet signifikante Wirkungen der Variablen "Geschlecht der Vpn", geht in seiner Diskussion aber nicht näher darauf ein. Da eine intensivere Berücksichtigung dieser differentiellpsychologischen Variablen die allgemeine Fragestellung der vorliegenden Arbeit überschreiten würde, soll "Geschlecht der Vpn" nur als Kontrollfaktor zur Verringerung der Fehlervarianz in den Varianzanalysen dienen.

5.2.2 Methode

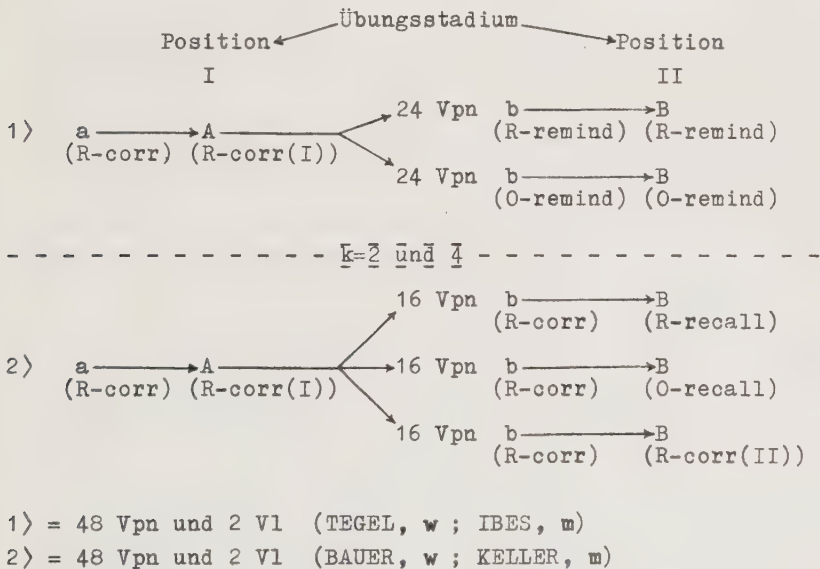
Als Besonderheit zu dem Experiment I ist zu vermerken, daß es im Rahmen zweier Semesterarbeiten unter Anleitung des Vf. durchgeführt wurde¹⁾. Der Methoden- und Ergebnisteil kann daher relativ kurz behandelt werden (zur Methode s. IBES u. KELLER, 1973; zu den Ergebnissen s. BAUER u. TEGEL, 1973. - Bei der jetzigen Darstellung werden für die einzelnen Versuchsbedingungen teilweise abweichende Kurzbezeichnungen gebraucht, die mnemotechnisch günstiger erscheinen und besser an die Diskussionen zum Experiment III angepaßt sind). Die Auswertung der Daten, die hier beschrieben werden soll, geht insofern über die in den Semesterarbeiten hinaus, als

1) Ich danke Fräulein Bauer, Herrn Ibes, Herrn Keller und Fräulein Tegel für die sorgfältige Durchführung der Versuche und die lebhaften Diskussionen.

sie sich mehr mit dem BM als mathematischem Modell auseinandersetzt, z.B. Probleme der Parameterschätzung und -invarianz behandelt; ferner wird die Prüfung der Hypothesen zu den Latenzzeiten eingehender diskutiert.

Versuchsplan

Der Versuchsplan läßt sich am einfachsten anhand folgender Übersichtsskizze erläutern. Die Skizze gibt den zeitlichen Verlauf einer Versuchssitzung für die fünf verschiedenen Testgruppen (R-remind, O-remind, R-recall, O-recall, R-corr (II)) wieder.



a und b bedeuten kurze Einübungslisten (je 16 Items); sie werden nicht ausgewertet. A und B sind die beiden Hauptlernlisten (je 40 Items). Die Vpn absolvieren die Listen nacheinander, also a...A...b...B, unter der jeweils in Klammern angegebenen Versuchsbedingung. Zu jeder Liste gehören drei Lerndurchgänge und ein Prüfdurchgang. Die Bedingungen R-corr (I)

und R-corr(II) unterscheiden sich nicht von R-corr, also der Standardbedingung; es sollen nur die unterschiedlichen Übungsstadien gekennzeichnet werden.

Nachdem alle Vpn (N=96) die Listen a und A unter Standardbedingungen absolviert hatten, wurden sie auf die fünf Testgruppen aufgeteilt. Dabei wurden den R-remind- und O-remind-Gruppen neue Instruktionen (s. unten: Versuchsdurchführung) vor Beginn der b-Liste gegeben, damit diese kurzen Listen zur Einübung der neuen "remind"-Bedingung dienen konnten. Für die drei anderen Testgruppen war eine neue Einübungsliste an sich nicht notwendig, sie hatten aber auch die b-Liste zu lernen, allerdings unter Standardbedingungen. Dieses Vorgehen hatte den Zweck, den zeitlichen Ablauf für alle Testbedingungen möglichst gleich zu halten, um so die fünf Gruppen in Position II besser vergleichbar zu machen. Auch das einheitliche Training aller Vpn mit Hilfe der Listen a und A kam dieser Absicht entgegen (s. S.89f.: h15).

Die R-recall- und O-recall-Gruppen erhielten neue Instruktionen erst vor dem letzten Prüfdurchgang, also nach den drei Lerndurchgängen zu Liste B. Die R-corr(II)-Gruppe bekam überhaupt keine neue Instruktion.

Die V1, d.h. die Bearbeiter der Semesterarbeiten, waren vom Vf. hinsichtlich der praktischen Versuchsdurchführung möglichst stark koordiniert worden.

Schließlich bedeutet $k=2$ und 4 im Übersichtsschema, daß der gesamte oben dargestellte Versuchsplan sowohl für $k=2$ als auch für $k=4$ durchgeführt wurde. Die angegebenen Vpn-Anzahlen sind also jeweils noch einmal zu halbieren, falls man sich für die Besetzung der $2 \times 5 = 10$ Testgruppen interessiert. - Die Faktoren: Geschlecht der Vpn, Geschlecht der V1 und Alternativenzahl k waren unter den Versuchsbedingungen in Position II gegeneinander ausbalanciert.

Innerhalb des Gesamtversuchsplanes sind noch einzelne Unterpläne herauszustellen in der Form, wie sie im allgemeinen Varianzanalysen zugrundeliegen (unabhängig davon, ob später

tatsächlich auch Varianzanalysen gerechnet wurden). Sie ergeben sich im wesentlichen schon aus der Beschreibung des Gesamtplanes und der Einzelbedingungen (R-remind, ...), sie sollen aber der Übersichtlichkeit halber und zur Erleichterung des Verständnisses der Versuchsdurchführung hier zusammengestellt werden.

R-corr(I): Faktor A: Geschlecht der Vpn (m, w)

Faktor B: Alternativenanzahl ($k=2$ und 4)

Faktor C: Rückmeldungen (r, f, b)

Es handelt sich also um einen $2 \times 2 \times 3$ -Plan mit abhängigen Messungen auf dem dritten Faktor. Es sind 96 Vpn auf die Zellen aufzuteilen.

Derselbe Plan gilt für die R- und O-recall-Bedingungen, zu je 16 Vpn.

R- und O-remind-Bedingungen, zu je 24 Vpn:

Da die remind-Listen "gemischt" sind, d.h. zur Hälfte "remind"-Items und zur Hälfte "R-corr"-Items enthalten, kommt zu den obigen Faktoren noch ein vierter (abhängiger) Faktor D hinzu:

Faktor D: "remind"-Items vs. "nonremind"-
("R-corr"-)Items

Es handelt sich also um einen $2 \times 2 \times 3 \times 2$ -Plan mit abhängigen Messungen auf den beiden letzten Faktoren.

Derselbe Plan gilt für den Vergleich R-corr(I) vs. R-corr(II) mit 16 Vpn; Faktor D wird dabei ersetzt durch

Faktor D': Übungsstadium (Position I vs. II).

Abhängige Variable für die R-corr- bzw. die R- und O-remind-Bedingungen ist: Anzahl der Wiederholungen der Antworten aus den Lerndurchgängen im Prüfdurchgang, getrennt für r-, f- und b-Items. Die relativen Wiederholungshäufigkeiten (pro Vp, Liste und Itemtyp ist durch 12 bzw. 6 zu dividieren) werden im Ergebnisteil mit rr, ff und bb bezeichnet.

Abhängige Variable für die R-recall- bzw. O-recall-Bedingung ist: Anzahl der korrekten Erinnerungen an die Antworten R bzw. Rückmeldungen O aus den Lerndurchgängen im Prüfdurchgang, getrennt für r-, f- und b-Items. Die relativen Erinnerungshäu-

figkeiten werden im Ergebnisteil mit rr', ff' und bb' bzw. mit rr'', ff'' und bb'' bezeichnet.

(Man vergegenwärtige sich am Rande, daß eine gute Gedächtnisleistung durch einen niedrigen ff-Wert und hohe ff'- bzw. ff''-Werte gekennzeichnet ist.)

Lernmaterial: Entsprechend BUCHWALDs Vorgehen dienten als Reizwörter möglichst einfach auszusprechende und bekannte deutsche Hauptwörter. Als Antwortalternativen wurden für die Bedingung k=2 die Ziffern 3 und 5 und für die Bedingung k=4 die Ziffern 3, 5, 7 und 9 gewählt. Die Listen a, A, b und B findet man im Anhang, S. 261; von jeder Liste ist die Permutation der Reizwortreihenfolge angegeben, die im jeweiligen 1. und 3. Lerndurchgang benutzt wurde. Für den 2. Lern- und den Prüfdurchgang fanden zwei andere Permutationen Gebrauch.

Die Hauptlisten A und B enthalten je 40 Items. Die jeweils ersten und letzten beiden Items werden später nicht ausgewertet, sie dienen nur zur Eliminierung von Anfangs- und Endeffekten. Die mittleren 36 Items bestehen aus 12 r-, 12 f- und 12 b-Items.

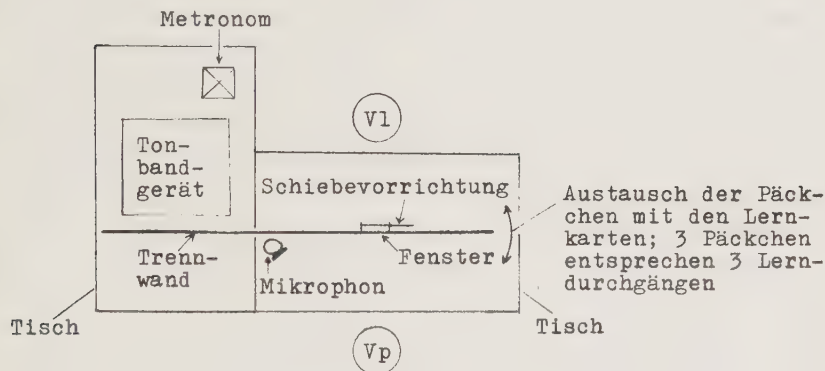
Die Antwortalternativen (Ziffern) und die Rückmeldungen r, f und b sind untereinander streng und bezüglich der mittleren 36 Positionen innerhalb der Listenpermutationen möglichst ausbalanciert. In den Permutationen der Liste B, die u.a. auch unter den "gemischten" R- bzw. O-remind-Bedingungen benutzt werden, wird diese Ausbalancierung auch noch auf den Item-Typ: R-corr-Item vs. R- bzw. O-remind-Item, ausgedehnt (für die Vpn wird natürlich erst im Prüfdurchgang ersichtlich, von welchem Typ die einzelnen Items sind).

Die Einübungslisten a und b sind analog aufgebaut.

Versuchsdurchführung

Vpn waren Studierende der Psychologie (m u. w; 1. bis 5. Sem.), die zur Erlangung ihrer Pflicht-Vp-Stunden an den Versuchen teilnahmen.

Die Versuchsanordnung sei durch folgende Skizze veranschaulicht:



Die Versuchsdurchführung für eine einzelne Vp verlief folgendermaßen:

Nach Aufnahme der Personalien hatte die Vp die Hauptinstruktion (s. S. 77-80, bzw. die in der k=2-Version) still zu lesen und anschließend laut in den wesentlichen Zügen zu wiederholen. Im Anschluß an das Training des manuellen Wechsels der Kärtchen erhielt die Vp die Liste a zur Einübung. Nach eventueller Klärung noch einiger Fragen der Vp wurden die drei Lerndurchgänge und der Prüfdurchgang mit der Liste A durchgeführt.

Für die R-corr(II)-Gruppe änderte sich nichts mehr bei den Listen b und B. Allgemein waren zwischen Liste A und Liste b ca. 5 min. Pause, zwischen Liste b und Liste B ca. 2 min. Pause vorgesehen.

Unter den R- und O-recall-Bedingungen gab es eine Abänderung des Verlaufs: Nach den drei Lerndurchgängen zur Liste B wurde vor dem Prüfdurchgang eine neue Instruktion gegeben (s. Anhang, S. 263).

Bei den R- und O-remind-Gruppen wurden vor Beginn der Liste b

neue Instruktionen gegeben, die sich dann bis zum Schluß nicht mehr änderten (s. Anhang, S. 263 und 264).

Während des Prüfdurchgangs schob der V1 die Kärtchen mit den Reizwörtern mit Hilfe einer Schiebevorrichtung in das Blickfeld der Vp. Das dabei erzeugte Geräusch (Anschlag) wurde auf dem Tonband registriert, ebenso die anschließend erfolgende Antwort der Vp. Die Zeiten zwischen diesen beiden Ereignissen mußten später noch mit Hilfe einer Stoppuhr gemessen werden.

Im Anschluß an den Prüfdurchgang für Liste B wurde die Vp dem postexperimentellen Interview unterzogen (s. S. 110f.).

Der gesamte Versuch dauerte ca. 1 1/2 Stunden.

Zur statistischen Auswertung

Es sei an den letzten Absatz in Kapitel 4 angeschlossen.

Bei vorliegender Arbeit handelte es sich nicht um eine Untersuchung, in der es um die Verwerfung oder Akzeptierung einiger weniger Hypothesen im Sinne eines "experimentum crucis" geht. Für eine solche Problemstellung ist strikte Sauberkeit in der Anwendung inferenzstatistischer Methoden zu fordern.

In unserem Fall dagegen sind ein mathematisches Modell, das BM, und ein der mathematischen Modellierung möglichst nahestehendes Hypothesensystem, nämlich die Latenzzeithypothesen H1 bis H9, auf "Anpassungsgüte" zu überprüfen. "... diese Überprüfung ist kein Akzeptierungs-Verwerfungs-Problem im Sinne der Inferenzstatistik. Man sollte die Anpassungsgüte nicht testen, sondern messen..." (nach BUSH, 1963, S. 432; s. auch HILGARD u. BOWER, 1966, S. 464).

Es kommt also nicht darauf an, für jede einzelne der bisher aufgestellten Hypothesen den adäquaten Test zu finden und durchzuführen. Vielmehr soll die Anpassungsgüte insofern gemessen werden, als das Muster der theoretischen Vorhersagen mit dem Muster der entsprechenden empirischen Daten auf deskriptiver Ebene miteinander verglichen werden soll (s. dazu z.B. HILGARD u. BOWER, 1966, S. 457, Tabelle 12-5).

Dieses Vergleichen enthält durchaus eine subjektive Komponen-

te. Es scheint aber vorzuziehen und auch vertretbarer zu sein, als eine größere Zahl von Signifikanztests zu rechnen, die teilweise auch noch an denselben Daten durchgeführt werden müßten. Die Anzahl und Abhängigkeit der Tests führen nämlich zu Veränderungen des nominalen Signifikanzniveaus, die nur in einfachsten Fällen noch zu beherrschen sind (s. dazu z.B. SACHS, 1969, S. 115 u. S. 495f.; HAYS, 1963, S. 576f.). Bei der Gesamtbeurteilung der unterschiedlichen Testausgänge spielte dann doch wieder eine subjektive Komponente mit hinein.

Als Folgerung aus diesen Überlegungen resultiert: Die Anzahl der inferenzstatistischen Tests soll möglichst klein gehalten werden. Nur die Hypothesen sollen statistisch abgesichert werden, die nach Meinung des Vf. besonders relevant sind und bei denen ein wirkliches Bedürfnis nach Entscheidungshilfe besteht. (Es spielen also auch Nutzensbetrachtungen und ferner zeitökonomische Gründe eine Rolle.) Dieses Vorgehen gilt für alle Experimente in der vorliegenden Arbeit.

Speziell im Experiment I sind zwei Hypothesen zu testen (s. S. 88f.):

h11: Test der Unabhängigkeitsannahme U1 unter der R-recall-Bedingung ($k=2$ und 4)

h13: Test der Unabhängigkeitsannahme U2 unter der O-recall-Bedingung ($k=2$ und 4).

Es sollen hier wohlgedenkt Nullhypothesen "bewiesen" werden, und das ist bekanntlich problematisch. Etwaige nichtsignifikante Testausgänge sind also entsprechend vorsichtig zu interpretieren.

Die Überprüfung obiger Hypothesen wurde mit Hilfe von $2 \times 2 \times 3$ -Varianzanalysen mit abhängigen Messungen auf dem dritten Faktor durchgeführt (s.S.93;s. dazu WINER, 1970, S. 330).

Die Kontrolle der Voraussetzungen für Anwendung dieser Varianzanalysen erfolgte nach Augenschein, da es für das exakte Vorgehen, nämlich Prüfung auf multivariate Normalverteilung und auf Homogenität der Varianz-Kovarianz-Matrix (WINER, 1970, S. 106 u. 123), keine Rechenprogramme gab. Im übrigen

sind Varianzanalysen relativ unempfindlich gegenüber Abweichungen von ihren Voraussetzungen (s. dazu BOX, 1953).

Bei den übrigen Vorhersagen, die auf deskriptiver Ebene abgehandelt werden sollen, ist nur noch auf die Auswertung der Hypothese h_3 (s. S. 87) einzugehen. Diese erfordert, aus den Gleichungen (1a) bis (1c) die Parameter c , u und g getrennt nach $k=2$ und $k=4$ zu schätzen.

Zunächst sind für die theoretischen Wiederholungswahrscheinlichkeiten $P(A_i/r)$, $P(A_i/f)$ und $P(A_i/b)$ ihre Maximum-Likelihood-Schätzungen rr , ff und bb einzusetzen (rr , ff und bb sind ja relative Wiederholungshäufigkeiten; s. dazu KREYSZIG, 1965, S. 179f.-Maximum-Likelihood-Schätzfunktionen haben im allgemeinen sehr günstige Eigenschaften, s. z.B. ATKINSON et al., 1965, S. 373-379).

Man erhält das Gleichungssystem (1a)' bis (1c)':

$$rr = \hat{c} \cdot (1 - \hat{u}) + \hat{c} \cdot \hat{u} \cdot \hat{g} + (1 - \hat{c})/k \quad (1a)'$$

$$ff = \hat{c} \cdot \hat{u} \cdot \hat{g} + (1 - \hat{c})/k \quad (1b)'$$

$$bb = \hat{c} \cdot \hat{g} + (1 - \hat{c})/k \quad (1c)'$$

Die Schätzfunktionen für c , u und g sind üblicherweise durch " $\hat{}$ " gekennzeichnet. Man hat also drei Gleichungen mit drei Unbekannten ($\hat{c}$, $\hat{u}$, $\hat{g}$), die sich eindeutig nach $\hat{c}$, $\hat{u}$ und $\hat{g}$ auflösen lassen:

$$\hat{c} = \langle (bb - 1/k)(rr - ff) \rangle / \langle bb - ff - 1/k(rr - ff) \rangle \quad (1a'')$$

$$\hat{u} = \langle 1/k(rr - ff) + ff - 1/k \rangle / \langle bb - 1/k \rangle \quad (1b'')$$

$$\hat{g} = \langle bb - ff \rangle / \langle rr - ff \rangle \quad (1c'')$$

Dieses Vorgehen beim Schätzen der Parameter entspricht dem BUCHWALDS im Zusammenhang mit der Prüfung der Gleichung (2a) mit Hilfe von (2b) und (2c) (s. S. 35). Ob $\hat{c}$, $\hat{u}$ und $\hat{g}$ die gleichen günstigen Schätzeigenschaften wie rr , ff und bb haben, ist nicht ganz klar.

Es wird nun die statistische Auswertung der Latenzzeiten behandelt. Die Zeiten wurden für alle Listen und Bedingungen gemessen, werden aber nur unter der R-corr(I)-Bedingung hier auch ausgewertet und dargestellt. Da die R-corr(I)-Gruppe die

Zeitdaten von 96 Vpn, je 48 Vpn für $k=2$ und $k=4$, erbrachte, war hier am ehesten eine gewisse Stabilität der Beziehungen zwischen den Medianen zu erwarten.

Gewöhnlich wurden die Zeiten im Anschluß an die einzelnen Versuchssitzungen des Tages von je zwei Vln von den Tonbandaufzeichnungen mit Hilfe einer Stoppuhr (Ablesegenauigkeit 1/100 sec) abgestoppt: Ein V1 bediente die Stoppuhr, der andere V1 trug die angesagten Zeiten in die Protokollbögen ein.

Dieses ist natürlich ein sehr grobes Verfahren (aus einer Testserie von 36 Antwortzeiten und fünfmaligem Abstoppen dieser Zeiten ergab sich als 95%-Konfidenzintervall für den "wahren" Einzelmeßwert: $t_w = t_{\text{beob}} \pm 0,07$ sec. - Erst in den Hauptuntersuchungen stand eine objektivere und genauere elektronische Zeitmeßapparatur zur Verfügung). Wie gleich zu erläutern ist, fällt die durch diese ungenaue Methode vergrößerte Fehlervarianz allerdings für die Auswertung kaum ins Gewicht.

Als Testverfahren wurden nämlich keine Tests für Mittelwertvergleiche, wie z.B. t-Tests oder Varianzanalysen verwendet, in denen eben die Fehlervarianz die bekannt große Rolle spielt. Der Verzicht auf diese Verfahren war durch folgende Auswertungsprozedur bedingt:

Getrennt für $k=2$ bzw. 4 und für alle sechs Antwortbedingungen ($rr, r\bar{r}, ff, f\bar{f}, bb, b\bar{b}$) wurden die Zeiten über alle 96 Vpn zusammengefaßt und daraus die zugehörigen, also $2 \times 6 = 12$, Mediane $T_{rr}, \dots, T_{b\bar{b}}$ berechnet. Nur diese Mediane werden zur Beurteilung der Hypothesen H_1 bis H_9 herangezogen.

Was nun die entsprechenden Verteilungen anbetrifft, deren Form und Typ doch im allgemeinen gerade von besonderem Interesse ist, so tragen die einzelnen Vpn zu jeder von ihnen mit mehreren und unterschiedlich vielen Werten bei. D.h., die Verteilungen sind weder vollständig unabhängig noch in kontrollierbarer Weise voneinander abhängig. Jedenfalls können die gewöhnlichen Tests auf Mittelwertsunterschiede nicht angewandt werden. - Die bei mathematischen Lernmodellen oft angenommene Homogenität von "subject-item-sequences" ist, auf

das vorliegende Problem angewandt, noch nicht hinreichend erprobt; s. dazu ATKINSON et al. (1965, S. 88).

Eine weitere Folgerung ist: Der Verteilungstyp, z.B. Gamma-Verteilung usw., kann nicht auf konventionelle Weise, z.B. mit Hilfe des χ^2 -Tests, überprüft werden, da dazu unabhängige Einzelmeßwerte Voraussetzung wären.

Erlaubt ist nur eine Inspektion nach Augenschein, indem man annimmt, daß die Verteilungen als Überlagerung von individuellen Einzelverteilungen deren Typ im großen und ganzen angemessen wiedergeben.

Man kann diese Überlagerung dadurch ein wenig mehr unter Kontrolle bringen, daß die Einzelverteilungen auf gleichen Mittelwert und eventuell auch gleiche Standardabweichung transformiert werden (z.B. WOODWORTH u. SCHLOSBERG, 1954, S. 264; ähnlich auch bei BUSH u. MOSTELLER, 1955, S. 319-327). Eigene Versuche in dieser Richtung brachten allerdings keinen augenscheinlichen Gewinn.

Die theoretische Analyse der Individualverteilungen selbst scheitert nun wiederum daran, daß diese Verteilungen maximal nur zwölf Werte enthalten können. Der einzige einschlägige "Kleine-Stichproben"-Test, der dem Vf. bekannt ist, bezieht sich auf Nachprüfung des Exponentialverteilungstyps (s. HOLLANDER u. WOLFE, 1973, S. 246-251). Aber die Annahme dieses Typs ist weder durch die tatsächlich beobachtete Form noch auch theoretisch aus den B-Diagrammen zu rechtfertigen.

Die oben beschriebene unkontrollierbare Abhängigkeit zwischen den sechs Verteilungen könnte nun dadurch aufgehoben werden, daß man von jeder Vp nicht alle Einzelmeßwerte getrennt heranzieht, sondern für jede Vp die sechs Mediane (T_{rr} , ..., T_{b5}) berechnet und diese als Ausgangsdaten behandelt. Es ergäben sich dann sechs in eindeutiger Weise abhängige Verteilungen, und es könnte z.B. eine doppelte Varianzanalyse oder der FRIEDMAN-Test durchgeführt werden.

Als Nachteile bleiben, daß jetzt die unterschiedlichen Beiträge der Vpn zu den Verteilungen nur verdeckt sind und daß die Struktur der Individualverteilungen, die, mindestens doch

der theoretischen Erwartung nach, die unterschiedlichen "Antwortwege" in den B-Diagrammen widerspiegeln sollte, der Analyse entzogen ist.

Es wird daher vorläufig für die Auswertung des Experiments I nur das Muster der experimentell gefundenen Beziehungen zwischen den Medianen (über alle Vpn) mit dem Muster des Hypothesenkomplexes H1 bis H9 verglichen. In der Gesamtauswertung nach Abschluß aller Experimente (s. 7.3) werden diese Mediane noch einmal in ein anderes Auswertungsverfahren einbezogen werden.

5.2.3 Ergebnisse

Die Darstellung der Ergebnisse folgt der Aufstellung der Hypothesen in 5.2.1.

R-corr: Hypothesen zu den Gleichungen (1a) bis (1c)

Die unter den R-corr-Bedingungen gefundenen relativen Wiederholungshäufigkeiten sind in Tab. 5.2.3-1 dargestellt. Man beachte: Zur R-corr(I)-Bedingung gehören 96 Vpn, zur R-corr(II)-Bedingung 16 Vpn (s. Versuchsplan auf S. 91).

Tab. 5.2.3-1: Relative Wiederholungshäufigkeiten rr, bb und ff (in %); R-corr-Bedingung in den Positionen I und II (Experiment I)

		rr	bb	ff
R-corr(I)	k=2	59,6	50,4	43,4
	k=4	47,3	27,4	27,8
R-corr(II)	k=2	68,7	60,4	30,2
	k=4	53,1	25,0	29,2

Diese Tabelle wird übersichtlicher durch folgende Profildarstellung (Abb. 5.2.3-1):

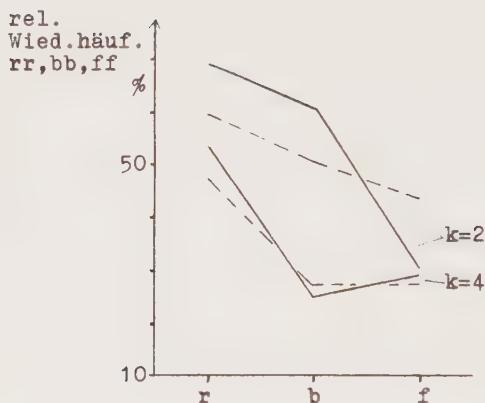


Abb. 5.2.3-1: Graphik zu Tab. 5.2.3-1; R-corr(I):----
R-corr(II):—

Im weiteren wird die Profildarstellung bevorzugt werden; für Parameterschätzungen stellen aber die Tabellen einen besseren Ausgangspunkt dar.

Zu h1: Erprobung der "Kärtchen"-Methode

Vergleicht man die Profile in Abb. 5.2.3-1 mit denen aus der BUCHWALDschen Untersuchung (1962) auf S. 24, so sind sie als hinreichend ähnlich zu beurteilen. "Unregelmäßigkeiten" wie z.B. der nicht streng monotone Abfall bei $k=4$ sind, jedenfalls bei der Geringfügigkeit der Abweichung, als Zufallsschwankungen zu werten; auch bei BUCHWALD kommen solche Profile vor.

Zu h2: Übungs-(Positions-)Effekt

Die Profile unter R-corr(II) verlaufen steiler als unter R-corr(I). Es besteht also ein Positionseffekt, der besonders für $k=2$ stark ausgeprägt ist. Bei $k=4$ und den ff-Werten findet man einen geringfügigen negativen Übungseffekt; die praktische Wirkungslosigkeit der Übung bezüglich der Bedingung " $k=4$ und ff" findet sich übrigens auch in späteren Experimenten (III und IV).

Zu h3: Invarianz von c , u und g gegenüber $k=2$ und $k=4$
 Aus den Gleichungen (1a) bis (1c) auf S. 98 ergeben sich folgende Parameterschätzungen:

R-corr(I)	$k=2$	$\hat{c} = -0,05$	$\hat{u} = 3,70$	$\hat{g} = 0,43$
	$k=4$	$\hat{c} = -0,06$	$\hat{u} = 3,21$	$\hat{g} = -0,02$
R-corr(II)	$k=2$	$\hat{c} = 0,36$	$\hat{u} = -0,06$	$\hat{g} = 0,78$
	$k=4$	$\hat{c} = 0,00$	$\hat{u} = \infty$	$\hat{g} = -0,18$

Diese offensichtlich merkwürdigen Ergebnisse werden später gemeinsam mit den anderen "Invarianz"-Resultaten diskutiert.

R-remind: Hypothesen zu den Gleichungen (2a) bis (2c)

In Tab. 5.2.3-2 und Abb. 5.2.3-2 sind die relativen Wiederholungshäufigkeiten dargestellt, und zwar getrennt nach den "reinen" R-remind-Items und den R-corr-Items.

Tab. 5.2.3-2: Relative Wiederholungshäufigkeiten rr, bb und ff (in %); "gemischte" R-remind-Bedingung in Position II (Experiment I)

		rr	bb	ff
R-remind-Items	$k=2$	66,7	66,7	54,2
	$k=4$	69,6	47,2	29,7
R-corr-Items	$k=2$	51,4	55,6	34,7
	$k=4$	45,8	29,2	33,4

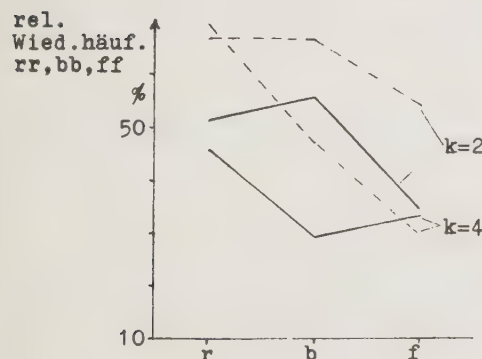


Abb. 5.2.3-2:

Graphik zu Tab. 5.2.3-2

R-corr-Items: ———

R-remind-Items: - - - -

Zu h4: Vergleich der R-corr- mit der R-remind-Bedingung
 Der Vergleich erfolgt über die geschätzten Differenzen
 (1a)-(1b) und (2a)-(2b). (Das %-Zeichen wird weggelassen.)
 $k=2$: (1a)-(1b)=51,4-34,7=16,7 > (2a)-(2b)=66,7-54,2=12,5;
 $k=4$: (1a)-(1b)=45,8-33,4=12,4 < (2a)-(2b)=69,6-29,7=39,8.

Der Vergleich unter $k=2$ spricht schwach gegen h4, der unter
 $k=4$ stark für h4.

Zu h5: Vorhersage von (2a) aus (2b) und (2c)

Aus (2c) folgt unmittelbar $bb=\hat{g}$. Also:

$k=2$: $\hat{g}=0,667$; $k=4$: $\hat{g}=0,472$.

Aus (2b) und (2c) ergibt sich:

$k=2$: $\hat{u}=0,542/0,667=0,812$; $k=4$: $\hat{u}=0,297/0,472=0,629$.

Eingesetzt in (2a) = $1-\hat{u} + \hat{u} \cdot \hat{g}$ folgt weiter:

$k=2$: $1-0,812+0,542=0,730$; $k=4$: $1-0,629+0,297=0,668$.

0,730 ist mit $rr = 0,667$ für $k=2$ zu vergleichen und 0,668 mit
 $rr = 0,696$ für $k=4$.

Die Übereinstimmung, verglichen mit BUCHWALDS Ergebnissen,
 ist mäßig.

Zu h6: Invarianz von u bzw. g gegenüber $k=2$ und $k=4$.

(s.o. unter h5)

Die Übereinstimmung ist wiederum sehr mäßig.

O-remind: Hypothesen zu den Gleichungen (3a) bis (3c)

In Tab. 5.2.3-3 und Abb. 5.2.3-3 sind die relativen Wieder-
 holungshäufigkeiten dargestellt, und zwar getrennt nach den
 "reinen" O-remind-Items und den R-corr-Items.

Tab. 5.2.3-3: Relative Wiederholungshäufigkeiten rr, bb und
 ff (in %); "gemischte" O-remind-Bedingung in
 Position II (Experiment I)

		rr	bb	ff
O-remind- Items	$k=2$	69,6	54,2	36,3
	$k=4$	68,0	45,8	29,7
R-corr- Items	$k=2$	62,4	57,0	30,5
	$k=4$	54,0	44,4	20,8

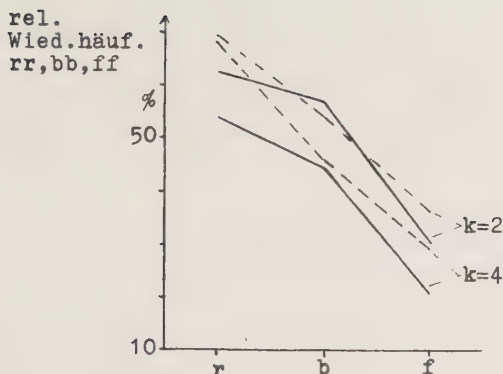


Abb. 5.2.3-3: Graphik zu Tab. 5.2.3-3; R-corr-Items: ———
O-remind-Items: -----

Zu h7: Vergleich der R-corr- mit der O-remind-Bedingung
Der Vergleich erfolgt über die geschätzten Differenzen
(1a)-(1b) und (3a)-(3b).

k=2: (1a)-(1b)=62,4-30,5=31,9 < (3a)-(3b)=69,6-36,3=33,3;

k=4: (1a)-(1b)=54,0-20,8=33,2 < (3a)-(3b)=68,0-29,7=38,3.

Die Vergleiche sprechen für h7. - Allerdings wird dieses für h7 positive Ergebnis dadurch relativiert, d.h. abgeschwächt, weil der von BUCHWALD besonders betonte Unterdrückungseffekt durch erinnertes (reminded) f nicht auftritt. Im Gegenteil, die Wiederholungen nach f unter der O-remind-Bedingung liegen sogar über denen unter der Standard-(R-corr-)Bedingung.

Zu h8: Vergleich von (1c) und (3c)

(1c) und (3c) sind durch bb zu schätzen.

k=2: (1c)=57,0 und (3c)=54,2; k=4: (1c)=44,4 und (3c)=45,8

Die Ergebnisse sprechen für h8.

Zu h9: Schätzung von c aus (3a) und (3b)

Die Schätzungen von c aus (3a) bekommen den Index r, die aus (3b) den Index f. Es ergibt sich durch Auflösen

k=2: $\hat{c}_r=0,38$ und $\hat{c}_f=0,28$; k=4: $\hat{c}_r=0,57$ und $\hat{c}_f=-0,12$.

Die Resultate sprechen gegen h9, d.h. also gegen die Unabhängigkeitsannahme U1 (s. S. 34), bei k=4 allerdings sehr viel stärker als bei k=2!

Der negative Wert von $\hat{c}_f$ bei k=4 resultiert daraus, daß nach Gleichung (3b) theoretisch $(1-c)/k < 1/k$ gilt; der entsprechende empirische Wert ist aber 0,297, also $> 1/k = 1/4 = 0,25$.

Zu h10: Invarianz von c gegenüber k=2 und k=4

Die Ergebnisse (s. unter h9) sprechen stark gegen h10.

R-recall: Hypothesen zu der Gleichung (4)

In Tab. 5.2.3-4 und Abb. 5.2.3-4 sind die relativen Erinnerungshäufigkeiten rr', ff' und bb' (s. S. 94) dargestellt. (Das strichlierte Profil gehört zu Experiment II und wird dort behandelt.)

Tab. 5.2.3-4: Relative Erinnerungshäufigkeiten rr', bb' und ff' (in %); R-recall-Bedingung in Position II (Experiment I)

	rr'	bb'	ff'
k=2	70,8	66,7	63,5
k=4	45,8	39,5	28,1

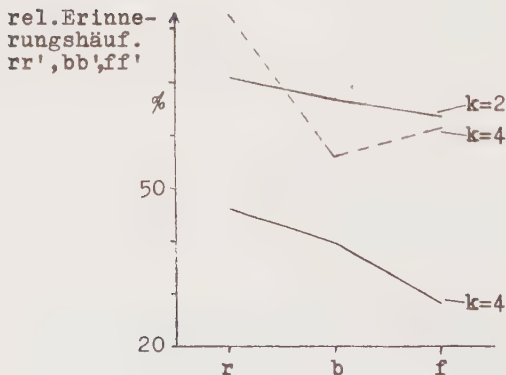


Abb. 5.2.3-4: Graphik zu Tab. 5.2.3-4; R-recall: — (R-recall(B) aus Experiment II mit k=4: ----)

Zu h11: Test auf Unabhängigkeit nach U1

Die hier interessierenden Ergebnisse der Varianzanalyse aus der R-recall-Bedingung sind (Signifikanzniveau $\alpha=5\%$): Signifikanz ($P=0,1\%$ und $P=2,4\%$) der Faktoren $k=2,4$ und $O=r,f,b$; Nichtsignifikanz ($P=55,1\%$) der Wechselwirkung zwischen diesen beiden Faktoren. Die Unabhängigkeitsannahme $U1:c_r=c_f=c_b=c$ ist also abzulehnen.

Trotz Nichtsignifikanz der Wechselwirkung $k \times O$ sind in Abb. 5.2.3-4 die Profile für $k=2$ und $k=4$ getrennt gezeichnet worden. Dem Trend nach ist U1 besonders für c_r und c_f unter $k=4$, weniger unter $k=2$, verletzt!

Anders als bei den Wiederholungsprüfungen unter R-corr spielt die Rückmeldung b bei den recall-Bedingungen eine untergeordnete Rolle. Es werden daher keine multiplen Vergleiche, wie z.B. der DUNCAN-Test, angegeben.

Zu h12: Invarianz von c gegenüber $k=2$ und $k=4$

Trotz Signifikanz des Faktors O über r , f und b gemittelt, ergibt sich aus Gleichung (4) für

$k=2$: $\hat{c}=0,34$ und für $k=4$: $\hat{c}=0,17$.

Das Ergebnis spricht eher für eine reziproke Abhängigkeit von k als für Invarianz.

O-recall: Hypothesen zu der Gleichung (5)

In Tab. 5.2.3-5 und Abb. 5.2.3-5 sind die relativen Erinnerungshäufigkeiten rr'' , ff'' und bb'' (s. S.94) dargestellt.

Tab. 5.2.3-5: Relative Erinnerungshäufigkeiten rr'' , bb'' und ff'' (in %); O-recall-Bedingung in Position II (Experiment I)

	rr''	bb''	ff''
$k=2$	46,8	19,8	45,8
$k=4$	60,4	46,7	47,9

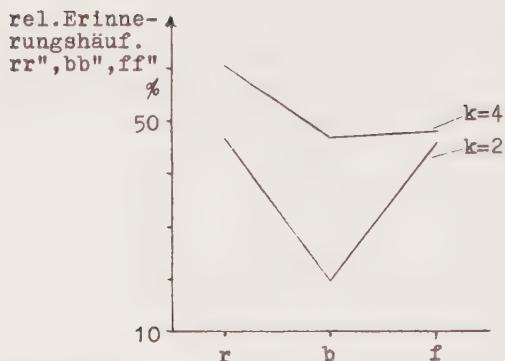


Abb. 5.2.3-5: Graphik zu Tab. 5.2.3-5

Zu h13: Test auf Unabhängigkeit nach U2

Die hier interessierenden Ergebnisse der Varianzanalyse aus der 0-recall-Bedingung sind (Signifikanzniveau $\alpha=5\%$): Nichtsignifikanz ($P=10,0\%$ und $P=6,5\%$) der Faktoren $k=2,4$ und $O=r,f,b$; Nichtsignifikanz ($P=54,1\%$) der Wechselwirkung $k \times O$. Die Unabhängigkeitsannahme U2: $1-u_r=1-u_f=1-u_b=1-u$ kann demnach beibehalten werden.

Trotz Nichtsignifikanz der Wechselwirkung $k \times O$ sind in Abb. 5.2.3-5 die Profile für $k=2$ und $k=4$ getrennt gezeichnet worden. Man sieht, daß für $k=2$ die Beziehung $1-u_r=1-u_f$ mit sehr guter Näherung erfüllt ist, für $k=4$ aber eher $1-u_r > 1-u_f$ zu gelten scheint!

Die Rückmeldung b spielt wiederum nur eine untergeordnete Rolle.

Zu h14: Invarianz von $1-u$ gegenüber $k=2$ und $k=4$

Aus Gleichung (5) ergibt sich gemittelt über r, f und b für $k=2$: $1-\hat{u}=0,07$ und für $k=4$: $1-\hat{u}=0,28$.

Der nicht übersehbare Unterschied ist aber trotzdem noch als nichtsignifikant ($P=10,0\%$; s.o. unter h13) zu beurteilen.

Zu h15: Parameterinvarianz zwischen den Gruppen

Stellt man die Parameterschätzungen aus den einzelnen Gruppen zusammen, so erhält man:

<u>R-remind</u>	k=2:	$\hat{g}=0,67$	$\hat{u}=0,81$	
	k=4:	$\hat{g}=0,47$	$\hat{u}=0,63$	
<u>O-remind</u>	k=2:	$\hat{g}=0,33$		
	k=4:	$\hat{g}=0,22$		(gemittelt aus $\hat{g}_r$ und $\hat{g}_f$ trotz $\hat{g}_f=-0,12$)
<u>R-recall</u>	k=2:	$\hat{g}=0,34$		
	k=4:	$\hat{g}=0,17$		
<u>O-recall</u>	k=2:	$\hat{u}=0,93$		
	k=4:	$\hat{u}=0,72$		

Zur Beurteilung der "Invarianz" s. 5.2.4.

Latenzzeiten: Hypothesen H1 bis H9

Die Mediane der Zeitverteilungen sind in Tab. 5.2.3-6 dargestellt.

Tab. 5.2.3-6: Zeitmediane (in sec) aus der R-corr(I)-Bedingung (Experiment I); T_k ist der Gesamtmedian aus $T_{rr} \dots T_{bb}$

	T_{rr}	$T_{r\bar{r}}$	T_{ff}	$T_{f\bar{f}}$	T_{bb}	$T_{b\bar{b}}$	T_k
k=2	2,22	2,44	2,53	2,30	2,45	2,55	2,45
k=4	2,05	2,63	2,65	2,45	2,45	2,54	2,55

Um das Beurteilen der Hypothesen zu erleichtern, werden die Mediane nach ihrer Rangfolge geordnet (s. auch Abb. 7.1.3-3 auf S. 201):

$$k=2: T_{rr} < T_{f\bar{f}} < T_{r\bar{r}} < T_{bb} < T_{ff} < T_{b\bar{b}}$$

$$k=4: T_{rr} < T_{f\bar{f}} < T_{b\bar{b}} < T_{bb} < T_{r\bar{r}} < T_{ff}$$

Hinter den folgenden Hypothesen ist kurz vermerkt, ob sie den empirischen Rangfolgen entsprechen (ja) oder nicht (nein).

H1: $T_{r\bar{r}} = T_{ff}$ nein für k=2, ja für k=4 (Hierbei wurden, zugegebenermaßen, die numerischen Unterschiede etwas subjektiv beurteilt.)

H2: $T_{rr} < T_{r\bar{r}}$ ja für k=2 und 4

H2': $T_{rr} < T_{ff}$ ja für k=2 und 4

- H3: $T_{rr} < T_{ff}$ ja für $k=2$ und 4
 H4: $T_{bb} < T_{rr}$ nein für $k=2$, ja für $k=4$
 H4': $T_{bb} < T_{ff}$ ja für $k=2$ und 4
 H5: $T_{rr} < T_{bb}$ ja für $k=2$ und 4
 H6: $T_{bb} < T_{bf}$ ja für $k=2$, nein für $k=4$
 H7: $(T_{ff} - T_{rr})_{k=2} < (T_{ff} - T_{rr})_{k=4}$: Es ist $2,30 - 2,22 = 0,08 < 2,45 - 2,05 = 0,40$, also ja
 H8: $(T_{bf} - T_{bb})_{k=2} < (T_{bf} - T_{bb})_{k=4}$: Es ist $2,55 - 2,45 = 0,10 > 2,54 - 2,55 = -0,01$, also nein
 H9: $T_{k=2} < T_{k=4}$ ja - Es gibt allerdings im einzelnen Umkehrungen, z.B. bei T_{rr} und T_{bf} .

Das Muster der empirischen Befunde entspricht befriedigend dem Hypothesenkomplex H1 bis H9. Die Güte der Entsprechung kann, grob, durch das Verhältnis ja/nein = 15/4 ausgedrückt werden. - Daß einige Hypothesen nur leicht verletzt und andere wiederum nur schwach bestätigt sind, wird im Augenblick vernachlässigt. Der Informationsverlust entspricht dem bei Gebrauch des Vorzeichen-Tests anstelle eines Rang- oder parametrischen Tests.

Postexperimentelles Interview: Es werden nur die für die weiteren Überlegungen besonders wichtigen Fragen angeführt. (Einer Reihe von Vpn konnten die Fragen aus Zeitmangel nicht gestellt werden.)

Frage 6: Fühltest du dich durch die Rückmeldungen "richtig" bzw. "falsch" in den Lerndurchgängen persönlich belohnt oder bestraft?

$k=2$ und 4: 1 ja, 75 nein

Frage 8: a) Versuchtest du, Reizwort-Zahl-Kombinationen mit der Rückmeldung "richtig" zuerst zu lernen...?

$k=2$: 11 ja, 30 nein $k=4$: 25 ja, 9 nein

b) ...und diese, wenn die Zeit dazu reichte, im

Stillen zu wiederholen?

k=2: 7 ja, 34 nein

k=4: 18 ja, 16 nein

Frage 12: Wenn im Prüfdurchgang ein Reizwort dargeboten wurde, kam es dann vor, daß dir

a) nur die Antwort aus den Lerndurchgängen einfiel?

k= 2 und 4: 54 ja, 18 nein

b) nur die Rückmeldung aus den Lerndurchgängen einfiel?

k= 2 und 4: 30 ja, 44 nein

c) beide zugleich einfielen?

k= 2 und 4: 63 ja, 10 nein

Frage 15:

a) Wenn du dich im Prüfdurchgang an die Antwort erinnerstest und an die Rückmeldung blank ("—"), hast du dann die erinnerte Antwort bevorzugt wiederholt?

k=2: 18 ja, 3 nein

k=4: 12 ja, 8 nein

b) Wenn du dich an die Antwort, aber überhaupt nicht an die Rückmeldung erinnerstest, hast du dann die erinnerte Antwort bevorzugt wiederholt?

k=2: 18 ja, 1 nein

k=4: 17 ja, 4 nein

Über a) und b) gemittelt wiederholten bevorzugt die erinnerte Antwort also 90% der k=2-Vpn und 71% der k=4-Vpn. - Inwiefern man bei der Beantwortung einer solchen Interviewfrage auch mit "Zufallsantworten" im Sinne von 1/k zu rechnen hat, ist nicht abzuschätzen.

Frage 20 (nur k=2): 30 Vpn gaben keine besondere Lernstrategie an; 12 Vpn sagten, sie hätten in den Lerndurchgängen bei der Rückmeldung f sofort versucht, sich die andere Alternativantwort einzuprägen.

Frage 1 aus den R- und O-remind-Gruppen: Empfindest du die Erinnerung an die Antwort (bzw. an die Rückmeldung) als Erleichterung?

k= 2 und 4: 26 ja, 14 nein

5.2.4 Diskussion

Es sollen zunächst die Ergebnisse besprochen werden, die für das weitere Vorgehen besonders wichtig sind. Die übrigen Resultate können dann relativ kurz behandelt werden.

Als erstes hat sich ergeben, daß sich die "Kärtchen"-Methode (im wesentlichen also "Beobachtungslernen") insofern bewährt hat, als die rr-bb-ff-Profile (s. Abb. 5.2.3-1) erwartungsgemäß ausgefallen sind. D.h., sie ähneln hinreichend denen BUCHWALDs (1962) und weisen den höheren Lernerfolg auf, der durch das Vorgeben von drei gleichartigen Lerndurchgängen beabsichtigt war.

Eng damit verbunden (s. S. 76) ist auch die Aufspaltung der Zeitmediane in eine, in gewissem Maße, vorhergesagte Ordnung.-Ein höheres Übungsniveau schlägt sich erwartungsgemäß in einer größeren Steilheit der Profile nieder.-Ein auf S. 82 vermuteter Vorteil des Beobachtungslernens, nämlich weitgehende Emotionsfreiheit, wird durch die eindeutige Beantwortung der Interviewfrage 6 gestützt.-Die allgemeine Regularität der Profile widerspricht auch nicht der Einführung des explizit ausgesprochenen "blank" anstelle "Schweigens des V1".

Von den Haupttests des BM spricht das Resultat aus der R-recall-Bedingung (h11) klar gegen die Unabhängigkeitsannahme U1: $c_r = c_f = c_b$. Die Unabhängigkeitsannahme U2: $1 - u_r = 1 - u_f = 1 - u_b$ braucht dagegen nicht verworfen zu werden; die Entscheidung ist allerdings knapp ($P=6,5\%$).

Betrachtet man die Abb. 5.2.3-4 und 5.2.3-5, so scheinen U1 und U2 für die k=2-Bedingung einigermaßen gut, für die k=4-Bedingung dagegen ziemlich schlecht erfüllt zu sein. In die gleiche Richtung geht die Folgerung aus den Schätzungen von c aus (3a) und (3b) (s. h9). D.h., die Befunde sprechen für eine asymmetrische Wirkung der Rückmeldungen zugunsten von r unter der k=4-Bedingung und für eine angenähert symmetrische Wirkung unter der k=2-Bedingung (s. auch Interviewfrage 8a und b, die ebenfalls auf eine besondere Aufmerksamkeitszuwendung auf r unter k=4 hinweist).

Die eindeutige Verwerfung von U₁, aber auch die Knappheit der Entscheidung bezüglich U₂, trifft das BM im Kern. Ist es also als falsifiziert zu betrachten? Dieses kann so nicht behauptet werden. Denn BUCHWALD (1969) hat unter der R-recall-Bedingung (bei ihm "apparent recall") Nichtsignifikanz des Faktors O=r,f,b erhalten. Eine Erklärung für diese diskrepanten Ergebnisse ist folgende:

Eine wesentliche Voraussetzung für die Gültigkeit der Unabhängigkeitsannahmen U₁ bis U₃ war die Verhinderung von Rehearsal. Dieses sollte im Experiment I durch relativ rasches Lesen der S-R-O-Sequenzen im Rhythmus eines Metronoms erreicht werden.

Vermutlich ist diese Absicht aber fehlgeschlagen, und zwar durch die Möglichkeit "visuellen" Rehearsals. Gemeint ist damit: Wenn die Vp die Lernkärtchen wechselte, sah sie bei jedem neuen Kärtchen sofort, ob darauf ein r-, f- oder b-Item stand. Sie konnte also, während sie die Reizwort-Zahl-Rückmeldungs-Sequenz aussprach, sich schon mit der Einprägung des Items beschäftigen.

Dabei hing eventuell die Intensität der Einprägungsbemühungen von der Rückmeldungsmodalität ab, sie konnte z.B. besonders stark bei r-Items unter der k=4-Bedingung sein (s. nochmals Interviewfrage 8a und b).

Die Begründung für diese asymmetrische Aufmerksamkeitszuwendung wäre dann einfach in der Informationshypothese und Strategieneklärung nach BUCHWALD (1962) zu sehen.

Die hier benutzte Instruktion (s. S. 77-80) legte der Vp nahe, "möglichst wenig Fehler zu machen". Das konnte sie ökonomisch erreichen, wenn sie sich unter der k=4-Bedingung auf die r-Items konzentrierte. Die Vp mußte ja subjektiv der Meinung sein, daß die Chancen, bei f-Items im Prüfdurchgang keinen Fehler zu machen, nur 1/3 waren. Daß es in Wirklichkeit bei den f-Items keine eigentlich richtigen Antworten gab, konnte die Vp nicht wissen.

So betrachtet, erscheinen das Ergebnis und seine Erklärung

nahezu trivial. Auf jeden Fall widerspricht eine von der Rückmeldungsmodalität abhängige Aufmerksamkeitszuwendung den Grundgedanken des BM. Forschungsstrategisch gesehen gibt es mindestens zwei Möglichkeiten, diese asymmetrische Aufmerksamkeitszuwendung zu verhindern, 1. durch erneute experimentelle Anstrengung, Rehearsal strikt zu unterbinden, 2. durch eine "symmetriebetonende" Instruktion.

Die erste Möglichkeit wird unter 6.2.1 und in den Experimenten III und IV weiterdiskutiert, die zweite Möglichkeit zunächst im folgenden Experiment II untersucht werden.

Vorerst soll aber in der Diskussion der Ergebnisse fortgefahren werden, und zwar mit dem Problem der Parameterschätzungen.

Die Resultate zu h_3 sind außerordentlich unangenehm. Da die Parameter c , u und g Wahrscheinlichkeiten darstellen, sind Schätzwerte $\hat{c}$, $\hat{u}$ und $\hat{g}$ kleiner als Null oder größer als Eins problematisch. Starke Abweichungen in diesen Richtungen sind überhaupt nicht zulässig, kleinere, z.B. $\hat{g} = -0,02$, können notfalls noch durch Zufallsschwankungen innerhalb der Profile erklärt werden.

Allerdings geben die sonstigen Ergebnisse (Profile und Parameterschätzungen aus den Spezialtestgruppen) keinen Hinweis darauf, daß man überhaupt extreme Parameterwerte, 0 oder 1, zu erwarten hätte, von denen aus kleinere "unzulässige" Abweichungen noch zu tolerieren wären.

Diese Diskrepanzen in den Parameterschätzwerten lassen sich so erklären:

Betrachtet man auf S. 98 die Auflösungen der Gleichungen (1a)' bis (1c)', also (1a)" bis (1c)", so sieht man, daß die Schätzfunktionen $\hat{c}$, $\hat{u}$ und $\hat{g}$ sehr empfindlich von den Profildaten rr , ff und bb abhängen. Es tauchen Differenzen dieser Werte in Zähler und Nenner auf; Zufallsschwankungen einer Differenz im Nenner um Null herum können erheblichen Einfluß auf den Gesamtwert des Bruches haben.

Hauptsächlich sind die Differenzen problematisch, in denen bb vorkommt, also $(bb-1/k)$ und $(bb-ff)$. In 6.2.3 werden die

Schwierigkeiten, die durch Schwankungen von bb zustandekommen, noch einmal direkt unter dem Gesichtspunkt der relativen Effizienz von r und f diskutiert.

Der Empfindlichkeit von $\hat{c}$, $\hat{u}$ und $\hat{g}$ in Abhängigkeit von rr , ff und bb entspricht in umgekehrter Richtung eine gewisse Unempfindlichkeit der theoretischen Wiederholungswahrscheinlichkeiten (S. 32) gegenüber Änderungen der Parameterwerte. Vertauscht man z.B. in (1b) die Werte von u und g , so ändert sich $P(A_1/f)$ überhaupt nicht. Dies spricht gegen eine problemlose Identifizierbarkeit der Parameter (s. auch das unten angeführte Beispiel).

Eine weitere Besonderheit unterscheidet das BM von anderen mathematischen Lernmodellen. Betrachtet seien im Augenblick nur die Standardexperimentalsituation (R-corr) und die zugehörigen Hauptgleichungen (1a) bis (1c) bzw. (1a)' bis (1c)'. Die Gleichungen enthalten drei Parameter (c, u, g), und ein Standardexperiment erbringt im wesentlichen drei empirische Daten (rr, ff, bb). Die sechs Größen, genauer: $\hat{c}$, $\hat{u}$, $\hat{g}$, rr , ff und bb , sind durch das Gleichungssystem (1a)' bis (1c)' stark eingeschränkt.

Im Gegensatz dazu sei ein anderes Lernmodell mit drei Parametern diskutiert, z.B. das "einfache Drei-Zustands-Modell" bei SUPPES et al. (1966, S. 100ff.; die Parameter sind c , b , q_1 . Von der Erweiterung auf Latenzzeiten sei im Augenblick abgesehen.). Die möglichen empirischen Daten überschreiten dort bei weitem die Anzahl drei, z.B. Verteilungen der Fehler im Durchgang n , der Versuche des letzten Fehlers, der Fehleranzahlen vor dem ersten Erfolg usw.

Die Parameterschätzungen $\hat{c}$, $\hat{b}$ und $\hat{q}_1$ sind nicht so stark mit den empirischen Daten verkettet, d.h. einzelne starke Zufallsabweichungen von den theoretisch zu erwartenden Werten können toleriert werden.

Um die starke Determiniertheit der zulässigen empirischen Häufigkeiten durch das BM, Gleichungen (1a)' bis (1c)', anschaulich zu machen, sei ein Beispiel angeführt:

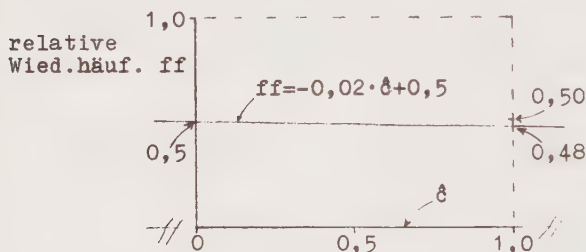
$k=2$; Gleichung (1b)': $ff = \hat{c} \cdot \hat{u} \cdot \hat{g} + (1 - \hat{c})/k$; ferner seien für $\hat{u}$

und $\hat{\epsilon}$ schon Werte bekannt: $\hat{u}=0,8$ und $\hat{\epsilon}=0,6$ (diese Werte sind einigermaßen repräsentativ für die bei BUCHWALD, 1963, und die hier unter der Spezialbedingung R-remind gefundenen Schätzungen, s. S. 109).

Die Gleichung (1b)' lautet damit: $ff = -0,02 \cdot \hat{\epsilon} + 0,5$.

Dieses ist jetzt also eine Bestimmungsgleichung für $\hat{\epsilon}$.

Graphisch sehen die Verhältnisse so aus:



Man liest leicht ab, daß ff nur Werte zwischen 0,48 und 0,50 (48% bis 50%) annehmen darf, falls ϵ nicht unzulässige Werte (<0 oder >1) zugeschrieben werden sollen. Es ist klar, daß eine derartige Einschränkung des Variationsbereiches von ff auf 2% empirisch leicht verletzt werden wird.

Nur am Rande seien noch die übrigen Ergebnisse, Häufigkeitsdaten betreffend, diskutiert.

Die Invarianz der Parameterschätzungen gegenüber $k=2$ bzw. $k=4$ ist nur mäßig, wenn überhaupt gegeben (s. h_6 , h_{10} , h_{12} , h_{14} . - h_3 wurde schon vorher abgehandelt). Wohlgermerkt spricht ein negativer Ausfall bezüglich der Parameterinvarianz nicht stark gegen das BM, die Unabhängigkeitsannahmen U_1 bis U_3 bleiben davon unberührt. Das BM wird nur etwas weniger ökonomisch hinsichtlich der Beschreibung oder Erklärung vieler Phänomene mit Hilfe weniger Parameter.

Die Resultate, die unter den R- bzw. O-remind-Bedingungen erhalten wurden, sind unbefriedigend.

Aus der Unregelmäßigkeit der R-corr-Profile in Abb. 5.2.3-2 auf S. 103 und dem fehlenden Unterdrückungseffekt von f unter der O-remind-Bedingung (s. Abb. 5.2.3-3 und die einschränken-

de Anmerkung zu h7 auf S. 105) kann geschlossen werden, daß in diesen Gruppen wegen der geringen Besetzung (12 Vpn pro Bedingung und Alternativenzahl) und eventuell auch infolge experimenteller Unzulänglichkeiten mit starken Zufallseffekten zu rechnen ist. Daher sind Ergebnisse, die den Hypothesen zu entsprechen scheinen, etwas fragwürdig.

Eine Beleuchtung etwaiger experimenteller Unzulänglichkeiten gibt die Interviewfrage 1 zu den remind-Bedingungen (S. 111): Von 40 Vpn empfanden 14 die Erinnerung an R bzw. O n i c h t als Erleichterung.

Unter diesen Gesichtspunkten ist auch die allgemeine Vergleichbarkeit der Spezialbedingungen belastet (s. zu h15 auf S. 108f.), obwohl das Muster der Parameterschätzungen nicht ganz ungeordnet scheint.

Auf eine gewisse Parallelität der Ergebnisse für Wiederholungstendenzen, ausgedrückt in $\hat{g}$, in Abhängigkeit von $k=2$ bzw. 4 sei hingewiesen: Sowohl aus der R-remind-Bedingung als auch aus der Interviewfrage 15 (S. 111) läßt sich eine stärkere Wiederholungstendenz unter $k=2$ als unter $k=4$ ablesen. Diese hier nur mit großer Vorsicht zu ziehende Schlußfolgerung soll später in Experiment III noch einmal mit "härteren" Daten untersucht werden.

Die Ergebnisse zu den Latenzzeiten waren schon am Anfang dieses Paragraphen als befriedigend beurteilt worden. Kleinere Abweichungen ergaben sich bei den Hypothesen $H1: T_{rr} = T_{ff}$ und $H4, H6, H8$ (diese Hypothesen betreffen Zeitmediane unter der Rückmeldung b).

Zu $H1$ kann im Augenblick nicht mehr gesagt werden; es müssen die Ergebnisse aus weiteren Experimenten abgewartet werden.

Die zweifelhaften Ausgänge bezüglich $H4, H6$ und $H8$ lassen sich vielleicht so erklären: Bei der Ableitung von $H4$ (S. 72) war vorausgesetzt worden, daß die Bedeutung von b für die Vpn sehr klar sei. Die Hauptinstruktion (s. S. 77-80) ist in dieser Hinsicht möglicherweise doch nicht so eindeutig.

In den Experimenten III bis V wird eine neue Instruktion

benutzt werden, in der die Rückmeldung b eingehender erklärt werden wird. Im unmittelbar nun folgenden Experiment II wurde allerdings noch keine Änderung in dieser Hinsicht vorgenommen, da hier ein anderer sehr viel wichtigerer Aspekt in den Vordergrund treten sollte.

Die Antworten der Vpn im postexperimentellen Interview sind teilweise in entsprechenden Zusammenhängen schon diskutiert worden. Es bleiben noch die Fragen 12 und 20.

Die Grundannahme BUCHWALDs, d.h. das Postulat zweier unabhängiger Gedächtnisprozesse (S-R und S-O), wird insofern durch die Antworten zu den Fragen 12a und 12b gestützt, als Erinnerungen nur an R oder nur an O tatsächlich vorzukommen scheinen (s. dazu S. 27f.). Die Unabhängigkeit der beiden Erinnerungen im Sinne der Annahme U3 (s. S. 34) wird aber durch die Beantwortung von 12c in Zweifel gezogen; danach scheinen R und O hauptsächlich gemeinsam erinnert zu werden.

In 7.1.2 wird ein objektiveres Maß für diese Abhängigkeit diskutiert werden.

Zu Frage 20 (nur $k=2$) gaben 12 von 42 Vpn an, sie hätten während der Lerndurchgänge bei der Rückmeldung f versucht, auf die andere Antwortalternative umzuschalten. Diese Strategie wirft einige Probleme auf, sowohl bezüglich der Behandlung durch das BM als auch hinsichtlich der Ableitung der Latenzzeithypothesen H3 und H7 (s. S. 71f.). In Experiment V wird auf diese Umschaltstrategie speziell eingegangen werden.

Kritische Anmerkungen zum experimentellen Vorgehen in Experiment I:

Die positiven Aspekte sind schon besprochen worden. Als negativ zu bewerten ist: Das (vermutlich) nicht verhinderte Rehearsal (s. S. 113); die Kompliziertheit der Instruktion und der Versuchssituation; die Schwierigkeit der manuellen Handhabung der Lernkärtchen, wodurch ein Teil des durch die drei Lerndurchgänge erzielten Lerngewinns wieder verloren ging; die im Laufe der Versuchssitzung abnehmende Motivation der Vpn, die z.B. durch die Monotonie beim dreimaligen Durchgehen

einer Lernliste als "beobachtender Lerner" bedingt sein konnte. - Es wird versucht werden, in den folgenden Experimenten diese negativen Punkte zu vermeiden, zumindest abzumildern.

Zusammenfassende Schlußfolgerungen: Es kann sich hier kurz gefaßt werden, da die Diskussion in größerem Rahmen in Kapitel 6 fortgesetzt wird.

Die Verwerfung der Unabhängigkeitsannahme U1 und das mindestens für $k=4$ nur knapp positive Ergebnis bezüglich U2 sprechen gegen das BM, jedenfalls im Rahmen des Experiments I. Die zwei vorgeschlagenen Möglichkeiten zur Korrektur der Experimentalsituation auf das BM hin (1. strikte Verhinderung von Rehearsal; 2. "symmetriebetonende" Instruktion) werden später noch diskutiert, führen aber, das sei vorweggenommen, nicht zum BM zurück.

Als Hauptargument gegen die Weiterverwendung des strikten BM im Rahmen der vorliegenden Arbeit sind die Schwierigkeiten bei der Parameterschätzung anzusehen, insbesondere was die Instabilität der Parameterschätzwerte anbetrifft. Der Vf. sah keine Möglichkeit, diese Probleme durch Modifikationen der BUCHWALDschen Axiome zu beheben: Durch jede Abänderung erhöhte sich entweder die Anzahl der Parameter, oder das BM verlor an überschaubarer Einfachheit, oder die spezifischen Grundannahmen mußten überhaupt aufgegeben werden.

Die Stabilität der Parameterschätzungen war aber auf S. 4 als Voraussetzung genannt worden, auf dem BM aufbauend ein mathematisches Latenzzeitmodell konstruieren zu können. Damit entfällt also die Weiterbehandlung der Fragestellung (2').

5.3 Experiment II: Verschärfte Instruktion und "speed"-Anweisung ($k=4$)

Dieses Ergänzungsexperiment wurde noch unter im wesentlichen gleichen Versuchsbedingungen wie das Experiment I durchgeführt, da zu dem damaligen Zeitpunkt der neue Versuchsraum

mit elektronischer Zeitmeßapparatur noch nicht zur Verfügung stand. Die oben in 5.2.4 angeführten negativen Seiten der experimentellen Situation wurden möglichst abgemildert.

5.3.1 Fragestellung und Hypothesen

In Experiment I war gefunden worden, daß unter $k=4$ -Bedingungen Antworten R und Rückmeldungen O besser erinnert wurden, wenn sie zu r-Items gehörten, O also gleich r war. Auch die Beantwortung der Interviewfrage 8 sprach für eine besondere Konzentration der Vpn auf die r-Items. Es wird vermutet, daß durch die Instruktion, "möglichst wenig Fehler zu machen", diese Einstellung bei den Vpn erzeugt wird, falls eine $k>2$ -Situation vorliegt, eben entsprechend der Informationshypothese (s. 2.3.2).

Die "asymmetrische" Einstellung sollte aufgehoben werden, wenn die Instruktion eine gleichmäßige Konzentration auf r- und f-Items betont: Der Informationswert von r und f soll also in dem Sinne gleich gemacht werden, auch in einer $k>2$ -Situation, daß bei r-Items die Antwort zu wiederholen und bei f-Items die Antwort nicht zu wiederholen ist.

Es wird folgende Hypothese aufgestellt:

Auch unter einer $k=4$ -Bedingung kann durch eine scharfe "symmetriebetonende" Instruktion Symmetrie in der Wirkung von r und f erzeugt werden. Speziell sollten in einem R-recall-Prüfdurchgang Antworten R aus r- und f-Items gleich gut behalten werden.

Es wurde in Experiment II noch eine Nebenfragestellung untersucht.

In Experiment I waren einige Vpn von der Kompliziertheit der Lernsituation, besonders auch von der manuellen Schwierigkeit beim Kärtchenwechseln im Rhythmus des Metronoms einigermaßen verwirrt und antworteten im Prüfdurchgang ziemlich nervös und schnell. Dabei war von den V1 versucht worden, die Versuchsatmosphäre möglichst entspannt zu halten und keinen Stress

und Zeitdruck aufkommen zu lassen. Nichtsdestoweniger verhielten sich also die oben erwähnten Vpn, als ob sie unter der Instruktion, möglichst schnell zu antworten, stünden, also unter einer "speed"-Anweisung.

Wie auf S. 52 angemerkt wurde, kann eine "speed"-Instruktion starken Einfluß haben, nicht nur trivialerweise eine größere Antwortschnelligkeit zu bewirken, sondern z.B. auch die Verhältnisse von Erfolgs- und Fehlerzeiten umzukehren. Es sollte also als Nebenfragestellung abgeschätzt werden, welchen Einfluß die "speed"-Einstellung einiger Vpn auf die Gültigkeit der Hypothesen H1 bis H6 haben könnte (H7 bis H9 entfallen, weil Experiment II nur unter $k=4$ durchgeführt wurde).

A priori besteht zunächst aber keine Veranlassung, die Hypothesen H1 bis H6 abzuändern. Sie werden also auch für das Experiment II aufgestellt.

5.3.2 Methode

In Experiment II spielte das eingehende Verständnis der Instruktion hinsichtlich der gleichen Aufmerksamkeitszuwendung gegenüber r und f eine entscheidende Rolle. Das Verständnis sollte dadurch erreicht werden, daß nur Vpn zum Versuch herangezogen wurden, die auch schon an der vorhergehenden Untersuchung teilgenommen hatten. (Zwischen den Durchführungen der Experimente I und II lagen mindestens zwei Monate.)

Die Hoffnung war, daß diese "trainierten" Vpn nicht mehr durch die Komplexität der Versuchsanordnung und durch die manuellen Schwierigkeiten beim Kärtchenwechseln so verwirrt wurden, daß darunter das Verständnis der Instruktion litt. Demzufolge waren diese Vpn in gewissem Maße "wissende" Vpn, was aber für die vorliegende Fragestellung kein Nachteil, eher ein Vorteil sein konnte.

Fernerhin wurden, um jeden das Verständnis verhindernden Stress abzubauen, statt drei wie in Experiment I, nun vier Lerndurchgänge erlaubt.

Versuchsplan

Der Versuchsaufbau entsprach im wesentlichen dem in Experiment I. Der Versuch wurde mit einer Versuchsgruppe unter der $k=4$ -Bedingung durchgeführt. Es nahmen 20 Vpn (15w, 5m) teil. Von diesen hatten in Experiment I 11 unter einer $k=4$ -Bedingung gelernt, 9 unter einer $k=2$ -Bedingung. -V1 war der Vf.

Die experimentellen Variablen waren 1. eine "symmetriebetonende" Instruktion; 2. eine "speed"-Instruktion.

Es wurden die Listen a, A und B (s. Experiment I) nacheinander gelernt. a hatte wieder die Funktion einer Einübungsliste. Die "speed"-Instruktion wurde erst in der B-Liste eingeführt.

Die Ergebnisse aus A (Profile und Latenzzeiten) sollten zur allgemeinen Abschätzung der Effekte aus "symmetriebetonender" Instruktion und den neu eingeführten vier Lerndurchgängen dienen. Insbesondere waren die Latenzzeiten aus der B-Liste unter der "speed"-Instruktion mit denen aus der A-Liste zu vergleichen.

Zum Abschluß der B-Liste wurde ein R-recall-Durchgang zur Prüfung der Haupthypothese bezüglich der Symmetrie von r und f durchgeführt.

Versuchsdurchführung

Der Versuchsablauf war (hinter den Prüfdurchgängen steht die jeweilige Kurzbezeichnung, die bei der Darstellung der Ergebnisse benutzt wird):

1. "symmetriebetonende" Instruktion (s. Anhang, S.264f.). An einem zusätzlichen Beispiel (LEIB 5 f) wies der V1 noch einmal explizit darauf hin, daß im Prüfdurchgang auf das Reizwort LEIB hin die Antwort 5 zu vermeiden sei.
2. a-Liste: vier Lerndurchgänge; Prüfdurchgang.
3. A-Liste: vier Lerndurchgänge; Prüfdurchgang (R-corr(A)).
4. Pause: ca. 5 min.

5. B-Liste: vier Lerndurchgänge; "speed"-Instruktion (mündlich durch V1):

"Es folgt jetzt der Prüfdurchgang. Dazu gibt es noch eine zusätzliche Anweisung. - Versuche, bitte, wie bisher, "richtig"-Alternativen zu wiederholen und "falsch"-Alternativen zu vermeiden, aber gleichzeitig auch so schnell wie möglich zu antworten."

Erster Prüfdurchgang zur B-Liste (R-corr(B; speed)).

Aufhebung der "speed"-Instruktion (mündlich durch V1):

"Der Prüfdurchgang unter der Schnelligkeitsbedingung ist nun beendet. Ich möchte dir nun noch einmal einen Lerndurchgang, d.h. den fünften, zu derselben Liste geben. Anschließend folgt wieder ein Prüfdurchgang, dieses Mal aber unter Normalbedingungen, d.h. ohne Schnelligkeitsanweisung."

Fünfter Lerndurchgang zur B-Liste.

R-recall-Instruktion (mündlich durch V1):

"Jetzt erfolgt der Prüfdurchgang, also ohne Schnelligkeitsbedingung. Deine Aufgabe dabei aber ist doch noch etwas abgeändert. Versuche, bitte, zu jedem Reizwort die Alternative zu sagen, die auf den Kärtchen bei diesem Reizwort gestanden hatte. Beachte also nicht mehr, ob hinter der Zahl "richtig", "falsch" oder "—" (blank) gestanden hatte. - Hast du dazu Fragen?"

Zweiter Prüfdurchgang zur B-Liste (R-recall (B)).

6. Postexperimentelles Interview

(Die Fragen des Interviews bezogen sich darauf, ob 1. die Vpn sich noch an das Lernmaterial aus dem Experiment I erinnert hatten und ob 2. die Vpn glaubten, sich gleichmäßig auf r- und f-Items konzentriert zu haben.)

Der Versuchsablauf dauerte ca. 70 min.

Die Auswertung der Daten sollte nur deskriptiv erfolgen.

3.3 Ergebnisse

Die relativen Wiederholungshäufigkeiten rr, bb und ff sind in der Abb. 5.3.3-1, die relativen Erinnerungshäufigkeiten rr', bb' und ff' in der Abb. 5.2.3-4 auf S. 106 dargestellt.

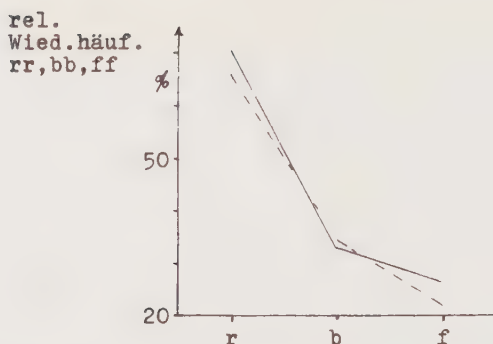


Abb. 5.3.3-1: Relative Wiederholungshäufigkeiten rr, bb und ff (in %); R-corr(A): —; R-corr(B; speed): ----; (Experiment II; k=4)

Die Mediane der Zeitverteilungen sind in Tab. 5.3.3-1 dargestellt.

Tab. 5.3.3-1: Zeitmediane (in sec) aus den R-corr(A)- und R-corr(B; speed)-Bedingungen (Experiment II; k=4). T_k ist der Gesamtmedian aus $T_{rr} \dots T_{b\bar{b}}$

	T_{rr}	$T_{r\bar{r}}$	T_{ff}	$T_{f\bar{f}}$	T_{bb}	$T_{b\bar{b}}$	T_k
R-corr(A)	2,31	3,35	3,31	3,37	3,40	3,45	3,36
R-corr(B; speed)	1,71	2,51	2,13	2,43	2,24	2,29	2,27

Um das Beurteilen der Hypothesen H1 bis H6 zu erleichtern, werden die Mediane nach ihrer Rangfolge geordnet (s. auch Abb. 7.1.3-3 auf S. 201):

$$\text{R-corr(A): } T_{rr} < T_{ff} < T_{r\bar{r}} < T_{f\bar{f}} < T_{bb} < T_{b\bar{b}}$$

$$\text{R-corr(B; speed): } T_{rr} < T_{ff} < T_{bb} < T_{b\bar{b}} < T_{f\bar{f}} < T_{r\bar{r}}$$

Die Hypothesen H1 bis H6 finden befriedigende Unterstützung. Unter R-corr(A) werden H4: $T_{bb} < T_{r\bar{r}}$ und H4': $T_{bb} < T_{ff}$ schwach verletzt; unter R-corr(B; speed) wird H4' schwach, H1: $T_{r\bar{r}} = T_{ff}$ stärker verletzt.

Das postexperimentelle Interview ergab:

Frage 1: Die Erinnerungen an das Lernmaterial in Experiment I wurden im Durchschnitt kleiner als 1% eingeschätzt. Nur eine Vp ($k=4$ in Experiment I) meinte, sich besonders an einige r-Items zu erinnern.

Frage 2: Im allgemeinen gaben die Vpn überwiegend an, sie hätten sich gleichmäßig auf r- und f-Items konzentriert. Der V1 zeigte ihnen darauf ihre individuellen Ergebnisse, besonders die R-recall(B)-Daten rr' und ff' (das Auszählen konnte innerhalb einer halben Minute vom V1 durchgeführt werden), und wies sie auf die offensichtliche Asymmetrie zugunsten von r hin. Die Vpn reagierten darauf leicht verlegen mit Achselzucken und meinten, sie könnten sich das nicht erklären oder irgendwie wäre es schließlich doch einfacher, r-Items zu lernen. Diese vom V1 unerwarteten Verlegenheitsreaktionen waren auffällig häufig und ausgeprägt.

5.3.4 Diskussion

Wie nicht anders zu erwarten (trainierte Vpn und vier Lerndurchgänge!), war das R-corr(A)-Profil steiler und auch "regulärer" als das R-corr(I)-Profil für $k=4$ in Abb 5.2.3-1 auf S. 102.

Befriedigend ist der mit der R-corr(A)-Bedingung vergleichbare Lernerfolg unter R-corr(B;speed). Die "speed"-Anweisung hatte also keinen sehr störenden Einfluß auf die Leistung im Prüfdurchgang. (Zum eventuell zu erwartenden Übungseffekt zwischen der A- und der B-Liste s. z.B. in Abb. 5.2.3-1 die nur wenig verschiedenen $k=4$ -Profile in den Positionen I und II.)

Ebenfalls befriedigend ist, daß auch unter der "speed"-Instruktion der Hypothesenkomplex H1 bis H6 unterstützt blieb. Der Unterschied zwischen "speed"- und "nonspeed"-Anweisung, also zwischen R-corr(B;speed) und R-corr(A), scheint hauptsächlich darin zu liegen, daß das Muster der Zeitmediane sich unter "speed"-Bedingungen zu kürzeren Zeiten verschiebt und

dabei etwas gestaucht wird.

Umgekehrt kann man vermuten, daß unter der Anweisung, sorgfältig und überlegt zu antworten, das Muster sich zu längeren Zeiten hin verschiebt und sich dabei auffächert. Dadurch könnten die Beziehungen zwischen den Medianen klarer zum Ausdruck kommen. Auf diesen Gedankengang wird in 6.3.4 zurückgekommen. Die Vermutung wird durch das Experiment III bestätigt werden.

Die Haupthypothese dieses Experimentes, Symmetrie von r und f unter "symmetriebetonender" Instruktion, ist augenscheinlich abzulehnen: Sowohl die Beurteilung der relativen Effizienz von r und f bezüglich b als auch die Asymmetrie der Erinnerungshäufigkeiten aus R-recall (B) ($rr'=82,9\%$; $ff'=61,2\%$) führen zu der Folgerung, daß r wirksamer war als f.

Daß bei der "Kärtchen"-Methode Rehearsal vermutlich nicht verhindert worden ist, wurde schon in der Diskussion zu Experiment I behandelt. Demnach kann unterschiedliches Rehearsal bei r- und f-Items entsprechend unterschiedlicher Einstellung der Vpn angenommen werden. Wie kommt es aber zu diesen unterschiedlichen Einstellungen bezüglich r und f trotz symmetriebetonender Instruktion?

Eine erste Erklärungsmöglichkeit ist die, daß die Vpn die Instruktion "rational" vollkommen verstanden haben. Entsprechend ihren ersten Antworten auf die zweite Interviewfrage schienen sie subjektiv überzeugt, sie hätten sich gleichmäßig auf r- und f-Items konzentriert.

Aus den anschließenden Verlegenheitsreaktionen kann vorsichtig geschlossen werden, daß die objektiv stattgefundene asymmetrische Behandlung von r- und f-Items eher unbewußt und automatisch vor sich gegangen ist.

Diese Schlußfolgerung hat offensichtlich mehr Ähnlichkeit mit dem Standpunkt THORNDIKES als mit der rein kognitiven Informationshypothese. Der Rest der vorliegenden Arbeit wird sich hauptsächlich mit dieser ersten Erklärungsmöglichkeit beschäftigen.

Eine zweite Erklärung wäre in der Darbietungsweise des Lernmaterials zu suchen. Dieses wurde wie beschrieben in vier (Beobachtungs-)Lerndurchgängen angeboten. Die Vpn sahen viermal z.B. das Item "GIFT 3 falsch". Es ist möglich, daß durch diese vier Durchgänge in den Vpn eine ähnliche Einstellung erzeugt wurde wie in den THORNDIKE-Experimenten, nämlich sich von Durchgang zu Durchgang verbessern zu wollen. Dieses ist aber, der Anlage des Lernmaterials nach, bei f-Items nicht möglich und kann daher irritierend zuungunsten von f wirken (s. dazu in einem verwandten Zusammenhang PARREREN, 1966, S. 81).

Ferner könnte eine auch nur geringfügige Unsicherheit im Verständnis der Instruktion zugunsten von r sich über die vier Lerndurchgänge potenziert fortgepflanzt haben.

Aus der Literatur sind keine Effekte bekannt, die in einer solchen, zugegebenermaßen sehr künstlichen Lernsituation auftreten können. Die Unsicherheit in der Interpretation der Ergebnisse wird zur Folge haben, daß in den weiteren Experimenten (III bis V) jeweils nur ein Lerndurchgang pro Liste durchgeführt wird. Da in diesen Experimenten bis auf einige Ausnahmen Rehearsal erlaubt sein wird, besteht die Hoffnung, daß auch in einem einzigen Lerndurchgang ein hinreichend hohes Lernniveau erzielt werden kann.

Eine dritte Erklärungsmöglichkeit schließlich ist die, daß die Vpn die Instruktion trotz zusätzlichen f-Beispiels in der Einübungsphase nicht genügend verstanden haben, daß sie also meinten, man müsse möglichst viele "Treffer" im Prüfdurchgang erzielen. Die letztere Einstellung würde bei $k=4$ eben eine Konzentration auf r-Items bewirken (Informationshypothese!). Dem Vf. (also dem Vl) scheint diese Erklärung relativ unwahrscheinlich. Weder aus den Bemerkungen der Vpn in der Einübungsphase noch aus der Beantwortung der zweiten Frage des postexperimentellen Interviews konnte entnommen werden, daß die Instruktion tatsächlich in diesem Sinne mißverstanden wurde.

Nichtsdestoweniger bedarf dieser Punkt noch einer Absicherung.

Im Hauptexperiment (III) wird eine Kontrollgruppe enthalten sein, die die Möglichkeit einer in obigem Sinne mißzuverstehenden Instruktion nahezu vollständig ausschließen wird.

Auch wenn die erste Erklärungsmöglichkeit vom Vf. favorisiert wird, so sind doch mindestens für das Experiment II die beiden anderen Erklärungsmöglichkeiten nicht auszuschließen. Möglicherweise liefert jede der drei Interpretationen einen gewissen Erklärungsanteil.

Zum Schluß noch zwei Bemerkungen.

Inspiziert man in Abb. 5.3.3-1 (S. 124) die Profile, drängt sich der Verdacht auf, daß es sich bei der Asymmetrie von r und f unter $k=4$ -Bedingungen um ein statistisches Artefakt handeln könnte, unabhängig von bzw. zusätzlich zu der Erklärung durch den g-Mechanismus.

Geht man von einem Zufallswiederholungsniveau von 25% aus, so kann rr um 75% bis 100% gesteigert werden, ff aber nur um 25% bis 0% gesenkt werden. Für die Asymmetrie könnte also ein "ceiling"-Effekt von unten her verantwortlich sein.

Der Verdacht wurde im vorliegenden Fall allerdings nicht bestätigt: Berechnung des Medians statt des arithmetischen Mittels der Wiederholungsskores nach r, b und f. - Nichtsdestoweniger scheint hier eine prinzipielle Schwierigkeit vorzuliegen.

Betrachtet man hieraufhin noch einmal das BM und schaltet die Erklärung durch den g-Mechanismus mit Hilfe der "delayed outcome"-Bedingung, also $u=0$, aus, so erhält man (s. S. 36):

$$k=2: \quad P(A_1/r) = 0,50 + 0,50 \cdot c \quad \text{und} \quad P(A_1/f) = 0,50 - 0,50 \cdot c$$

$$k=4: \quad P(A_1/r) = 0,25 + 0,75 \cdot c \quad \text{und} \quad P(A_1/f) = 0,25 - 0,25 \cdot c$$

Für jeden beliebigen Wert von c hat man also bezüglich $1/k$ für $k=2$ Symmetrie von r und f, für $k=4$ dagegen Asymmetrie zuungunsten von f. Diese durch das unterschiedliche Zufallsniveau bedingten Verhältnisse sind von BUCHWALD nicht diskutiert, möglicherweise auch nicht gesehen worden.

Es wird in 6.3.4 ein Maß für die relative Effizienz von r und f vorgeschlagen ($re(r!)$), das diese Schwierigkeit vermei-

det.

Auf die erste Frage des Interviews hatte nur eine Vp geantwortet, sie hätte sich noch an mehrere der r-Items ($k=4!$) aus Experiment I erinnert. Prinzipiell besteht die Möglichkeit, daß auch den anderen ehemaligen $k=4$ -Vpn jetzt in Experiment II besonders die r-Items vertraut waren, weil diese eben in der vorhergehenden Untersuchung aufmerksamer beachtet worden waren (s. Abb. 5.2.3-4 und 5.2.3-5; S. 106 u. 108). Die symmetriebetonende Instruktion hätte sich dann, wenn auch von den Vpn unbemerkt, unter Umständen gegen den Vertrautheitseffekt hinsichtlich der r-Items nicht durchsetzen können. Dieses wäre ein Einwand gegen das Vorgehen, "trainierte" Vpn aus Experiment I in das Experiment II zu übernehmen.

Zur Überprüfung dieses möglichen Einwandes wurde das R-recall (B)-Profil getrennt für die neun ehemaligen $k=2$ -Vpn, für die die obige Argumentation nicht zutrifft, berechnet und ebenso für die elf ehemaligen $k=4$ -Vpn. Es ergab sich kein nennenswerter Unterschied.

6. Zwischendiskussion

Entsprechend den Schlußfolgerungen am Ende von 5.2.4 wird im weiteren das BM als striktes mathematisches Modell nicht mehr zugrundegelegt. Es entfällt also die Weiterbehandlung der Fragestellung (2') zugunsten der Fragestellungen (4) und (5) (s. S. 5). Hieraus folgt eine Wiederaufnahme der allgemeinen Diskussion zur Informationshypothese, wie sie in Anschluß an BUCHWALD (1962) in 2.3.2 dargestellt wurde.

6.1 Folgerungen aus der Voruntersuchung

Es sollen hier kurz die Punkte aus den Diskussionen zu den Experimenten I und II (5.2.4 und 5.3.4) zusammengestellt werden, die für das weitere Vorgehen wichtig sind.

Von den Besonderheiten der "Kärtchen"-Methode sollen das "Beobachtungslernen" und das explizit ausgesprochene "blank" weiterverwendet werden. Die Durchführung mehrerer (3 bis 5) gleichartiger Lerndurchgänge erzeugt für die Interpretation Unsicherheit und wird daher zugunsten eines einzigen Lerndurchgangs fallengelassen.

Die Beschäftigung mit den Vor- und Nachteilen des Beobachtungslernens (s. S. 82) und die Durchsicht der einschlägigen Literatur (z.B. MARX u. WITTER, 1972) führten zu der Idee, daß die Gegenüberstellung von "aktivem" und "beobachtendem" Lernen nicht nur für die Trennung von Information und Effekt nützlich sein kann. Es erschien aussichtsreich, die beiden Lernmodalitäten zur Herstellung von symmetrischen und asymmetrischen Situationen heranzuziehen (s. Fragestellung (4)). Diese Idee wird in 6.3.3 weiter ausgeführt.

Die Hauptgründe für die Ablehnung des BM als Ausgangsbasis für ein Latenzzeitmodell sind:

1. Schwierigkeiten, die sich aus den Parameterschätzungen ergeben (s. S. 114ff. und 119).
2. Ablehnung des BM innerhalb des experimentellen Rahmens

von Experiment I wegen Verletzung der Unabhängigkeitsannahme U1. Dadurch wird aber das BM nicht allgemein falsifiziert, da in Experiment I die Verhinderung von Rehearsal offenbar nicht gegeben war. Es wird in 6.2.1 begründet werden, warum die forschungsstrategische Möglichkeit, die strikte experimentelle Verhinderung von Rehearsal zu erzwingen, nicht intensiver verfolgt wird.

Es wurde ferner erörtert, daß unterschiedliches Rehearsal zugunsten von r in $k=4$ -Aufgaben durch die Instruktion bewirkt wird (Informationshypothese). Die "symmetriebetonende" Instruktion in Experiment II ($k=4!$) konnte keine symmetrische Wirksamkeit von r und f erzeugen. (Eine Klärung bezüglich des Verständnisses der Instruktion ist für Experiment III vorgesehen.)

Falls also Rehearsal stattfinden kann, ist die Erklärung der Ergebnisse durch die Informationshypothese (BUCHWALD, 1962) adäquater als die durch das BM (1969). Allerdings wird durch die Resultate von Experiment II die rein kognitive Informationshypothese in Frage gestellt, jedenfalls für $k=4$ -Situationen, und eine mehr unbewußte automatische Informationsverarbeitung suggeriert. Hier deutet sich eine gewisse Rückbeziehung auf den THORNDIKESchen Standpunkt an.

Die Ergebnisse der Latenzzeitmessungen sind zufriedenstellend. Hinsichtlich der Komponentenanalyse der Zeitverteilungen besteht allerdings das Bedürfnis, die $(c, 1-u)$ -Wege von den GK-Anteilen separieren zu können.

6.2 Kritische Anmerkungen zu BUCHWALDs Arbeiten

Bevor auf einige problematische Punkte in BUCHWALDs Untersuchungen eingegangen wird, sollen seine Arbeiten und Beiträge zum Problem der relativen Effizienz von r und f allgemein gewürdigt werden. BUCHWALD hat die Informationserklärung zwar nicht als erster und einziger angeführt, sie aber besonders deutlich und klar ausgearbeitet. Man kann sein einfaches und

elegantes Modell (1969) mit den nahezu provozierenden Unabhängigkeitspostulaten U1 bis U3 als einen Extremfall des reinen Informationsstandpunktes betrachten, wodurch der Kontrast zum Verstärkungsprinzip THORNDIKES prägnant wird.

6.2.1 Notwendigkeit von Nonrehearsal?

Im Anschluß an den negativen Ausfall des Tests der Unabhängigkeitsannahme U1 war die Rolle des Rehearsal bzw. die Notwendigkeit der Verhinderung von Rehearsal in den Vordergrund gerückt worden. Es entstand der Entscheidungspunkt, entweder das BM in voller Strenge aufrechtzuerhalten zu suchen und die Forderung des Nonrehearsal mit möglichster Genauigkeit zu erfüllen oder das BM durch Abschwächung der theoretischen und experimentellen Forderungen zu modifizieren, dabei unter Verlust der mathematischen Eleganz. Es wurde sich gegen die erste Möglichkeit entschieden, und zwar aus folgenden Gründen:

1. Der Vf. hatte in den Experimenten I und II die Erfahrung gemacht, daß viele Vpn sich verärgert über die Schnelligkeit in den Lerndurchgängen äußerten, daß sie also nicht genug Zeit hätten, sich die Items gehörig einzuprägen.

Wie nun auch immer eine noch stärkere Verhinderung von Rehearsal experimentell durchgeführt würde, es ist jedenfalls mit noch stärkerer Verärgerung und größerem Stress bei den Vpn zu rechnen. Im Bemühen also, die Experimentalsituation mehr zu kontrollieren, würde der Vf gerade verstärkt unkontrollierbare Reaktionen in den Vpn evozieren. Solche Emotionen entsprechen sicherlich nicht dem kognitiv betonten BM. Es ist ferner zu erwarten, daß das Lernniveau sich bis in die Nähe des Zufallsniveaus absenkt. Diese Erwartung wurde durch die Ergebnisse des Experiments IV (s. 7.2.1) bestätigt.

2. LIU (1972, S. 368) hatte ebenfalls versucht, die Zeit für Rehearsal zu minimieren, und kommt zum Schluß, daß auch ein sehr kurzes Zeitintervall immer noch ausreicht, r-Items, nicht aber f-Items, innerlich zu wiederholen aufgrund eines "established habits", und daß das BM diese mögliche Bevorzu-

gung von r nicht genügend erfaßt. (Zu der Größenordnung von Rehearsalgeschwindigkeiten s. NORMAN, 1970, S. 179.)

3. Die Künstlichkeit eines Laborexperiments muß nicht unbedingt gegen seine Durchführung sprechen. Findet aber eine experimentelle Manipulation überhaupt keine Entsprechung im Alltag, so ist ihr dann doch mit Mißtrauen zu begegnen. Dem Vf. scheint es, daß ein Richtig-Falsch-Paradigma unter strikten Nonrehearsal-Bedingungen im wesentlichen zur Aufrechterhaltung des BM dienen würde, daß ihm aber sonst keine weitere psychologische Bedeutung zukäme. (Zum Problem der Künstlichkeit von Experimenten, durch die Zusammenhänge eher verdeckt oder verzerrt als aufgedeckt werden, s. REITMANN, 1970.)

4. Dieser letzten Literaturangabe kann noch eine Bemerkung von ESTES (1971, S. 23) hinzugefügt werden (s. auch HERRMANN u. STAPF, 1973). Danach gibt es einen Trend, zumindest im Bereich der Untersuchungen mit Erwachsenen, eher Regeln, Prinzipien und Strategien in Abhängigkeit von Belohnung und Bestrafung als Verhaltenseinheiten zu betrachten als elementarere S-R-Verbindungen. In einem strikten BM wäre dafür kein Raum.

Diese vier Gründe trugen mit mehr oder weniger Gewicht dazu bei, keine größeren Anstrengungen zur strikten Verhinderung von Rehearsal zu unternehmen und, daraus folgend, das BM als mathematisches Modell fallenzulassen. Dieser Entschluß wurde auch erleichtert durch die Abwägung der für das BM sprechenden Befunde (hauptsächlich in BUCHWALD, 1969) gegen die Einschränkungen, wie sie beiläufig bei der Darstellung des BM in dieser Arbeit gemacht wurden (z.B. Unsauberkeiten bei statistischen Tests, Nichtdurchführung von möglichen Tests und Parameterschätzungen, Versäumnis des Tests von U3).

Es fällt ferner auf, daß in der Literatur nach 1969 kein Bezug auf das BM als mathematisches Modell genommen wird, mit der schon erwähnten Ausnahme (s.S. 41f.). Aber auch HUMPHREYS (1972) gelingt es nicht, das Problem der relativen Effizienz

von r und f mit seinem Modell überzeugend in den Griff zu bekommen.

Es scheint, als ob die mathematische Modellierung einer Lernsituation unter der Nichtkorrektur-Bedingung doch erheblich schwierigere Probleme aufwirft als die schon gut bewährte Behandlung von PAL-Experimenten unter der Korrektur-Prozedur (s. auch auf S. 42f. NAHINSKY, 1964, 1966, 1967; usw.).

Nichtsdestoweniger wird das BM in modifizierter Form in dieser Arbeit verwendet (s. Kapitel 4 und 6.3.1). Dabei wird versucht, seine Struktur so weit wie möglich beizubehalten, um die Vorteile eines formalisierten Ansatzes nicht zu verlieren.

Trotz obiger Argumentationen gegen die Forderung strengen Nonrehearsals wird versucht werden, allerdings nur am Rande, Versuchsgruppen mitzuuntersuchen, in denen Rehearsal möglichst verhindert wird; teils um nachzuprüfen, ob die obige Begründung (1.) wirklich stichhaltig ist, teils um abzutesten, inwieweit sich bei Verminderung der Rehearsalmöglichkeiten die neue Methode des "Verbalisierens" bewähren würde, was die mindestens trendmäßige Annäherung an die Vorhersagen des strengen BM betrifft.

Zum Abschluß dieses Abschnitts einige einschlägige Literaturhinweise zur Rolle des Rehearsal beim Problem der relativen Effizienz von r und f: ESTES (1971, S. 29); ALLEN u. ESTES (1972); MEDIN (1972); LONGSTRETH (1970, 1971); LONGSTRETH u. BARGE (1970). Insbesondere LONGSTRETH diskutiert die unterschiedlichen Ergebnisse in Abhängigkeit davon, ob Rehearsal großzügig oder beschränkt erlaubt ist. Von allen genannten Autoren wird die "frequency theory" von EKSTRAND et al. (1966) positiv bewertet.

6.2.2 Einschränkungen der Informationshypothese

In diesem Abschnitt sollen mehrere Punkte übersichtlich zusammengestellt werden, die im Verlauf dieser Arbeit schon an verschiedenen Stellen angedeutet wurden. Insbesondere werden

einige Einschränkungen der rein kognitiven Informationshypothese und der Strategieneklärung behandelt (s. 2.3.2; BUCHWALD, 1962).

Es wurde schon auf S.26 (Punkt 4) bemerkt, daß, wenn keine näheren Angaben zum Rehearsal gemacht werden, ein gewisser Widerspruch zwischen der Informationserklärung (1962) und dem BM (1969) besteht. Falls Rehearsal erlaubt ist, kann die Informationserklärung angewandt werden; dann sind allerdings nicht die Unabhängigkeitsannahmen U1 bis U3 des BM zu garantieren. Falls Rehearsal nicht erlaubt ist, ist keine Gelegenheit für unterschiedliche Strategien aufgrund unterschiedlichen Informationsgehaltes von r und f gegeben, und es wird das BM möglich.

BUCHWALD macht nun allerdings diese Einschränkungen nicht explizit. Es ist sogar zu vermuten, daß er sie nicht gesehen hat. Denn in seiner Arbeit von 1962 (S. 72), in der er die Informationshypothese und Strategieneklärung vertritt, wird versucht, Rehearsal zu unterbinden ("...E attempted to minimize the time between S's response and the next succeeding stimulus presentation...").

Ein weiterer Punkt: Wie hängt die Informationserklärung von der Versuchssituation und der Instruktion ab? Betrachtet werden jetzt nur $k > 2$ -Experimente.

1. In den THORNDIKE-Experimenten waren mehrere Lern- (=Prüf-)Durchgänge vorgesehen. Hier ist die Anwendung der Informations- und Strategieneklärung plausibel und problemlos.

2. In den BUCHWALD-Experimenten unter der IT-Prozedur gibt es nur einen Lerndurchgang und nur einen Prüfdurchgang. Die Instruktion lautet, wie auf S. 22 angegeben, oder noch eindeutiger (BUCHWALD, 1967, S. 318f.): "... (Ss) should try to give as many correct answers as they could on this (test)trial." Danach werden sich die Vpn vernünftigerweise auf r-Items konzentrieren, um ihre Trefferzahl im Prüfdurchgang zu maximieren. Daß es bei f-Items keine eigentlichen Treffer gibt, können die Vpn nicht wissen.

Die Möglichkeit der Informations- und Strategieneklärung bleibt also erhalten. Der Rückmeldung f wird allerdings aufgrund dieser Instruktion, salopp ausgedrückt, praktisch keine Chance gegeben.

3. Ein BUCHWALD-Experiment (e i n Lern- und e i n Prüfungsdurchgang) kann aber auch noch anders behandelt werden. Dieses wurde schon in Experiment II angestrebt und soll in den weiteren Untersuchungen ausgearbeitet werden.

Da der Vp nur einmal Gelegenheit geboten wird, nämlich im Prüfungsdurchgang, zu zeigen, was sie im Lerndurchgang gelernt hat, ist es nicht unvernünftig zu fordern, daß sie diese Gelegenheit für f -Items genauso ausnutzt wie bei r -Items. Danach ist die Instruktion so zu gestalten, daß die Vp eine Maximierung des Lernerfolgs dadurch anstrebt, daß sie sich gleichermaßen um (Nicht-)Wiederholungen bei (f -) r -Items bemüht.

Eine rein kognitive Informationshypothese, die ein vollkommen bewußtes, rationales Strategieverhalten der Vp voraussetzt, würde dann, trotz $k > 2$, Symmetrie in der Wirkung von r und f voraussagen (s. dazu auch BLÖSCHL, 1969, S. 162).

Die rein kognitive Informationshypothese nach BUCHWALD (1962) wird hier also um den Aspekt der Instruktionsabhängigkeit erweitert. Sie ist in 6.3.3 noch einmal zu modifizieren; dabei wird das rein kognitiv gesteuerte Strategieverhalten der Vpn angezweifelt.

Die Punkte 1. bis 3. zusammenfassend soll nachdrücklich betont werden, wie sehr die Antwort auf die Frage nach der relativen Effizienz von r und f von der Experimentalsituation abhängen kann. Dabei wurde im Augenblick nur die $k > 2$ -Situation betrachtet. Davor auf die allgemeine Abhängigkeit vom Aufgabentyp ($k=2$; $k > 2$; Begriffsbildung) aufmerksam gemacht zu haben, ist ausdrücklich das Verdienst der "Informationstheoretiker", wie z.B. BUCHWALD.

Es wurde eben ein rationales Strategieverhalten angesprochen. Die Informationshypothese, wie sie im allgemeinen dargestellt wird, beinhaltet, daß die Vp die Information, die in den

Rückmeldungen steckt, bewußt verarbeitet und ihre Lernstrategie danach planvoll ausrichtet. Daß dieses in gleicherbewußter Weise für die $k=2$ -, $k>2$ - und Begriffsbildungsaufgaben geschieht, soll bezweifelt werden.

Nach dem hierarchischen Modell von GAGNÉ (1969, S. 51f.) stellt der zur Begriffslernaufgabe gehörende Lerntypus eine deutlich höhere Stufe dar als der zur $k>2$ -Aufgabe gehörende.

Daß die Begriffslernaufgabe als Problemsituation aufgefaßt wird und die Rückmeldungen im Sinne der Informationshypothese bewußt strategisch ausgenutzt werden, scheint problemlos zu sein.

Daß dieses auch für die $k>2$ -Aufgabe gilt, ist aufgrund der Ergebnisse von Experiment II anzuzweifeln: Hier wird eine nicht so voll bewußte, mehr automatische Informationsverarbeitung angenommen werden (s. 6.3.3).

Die $k=2$ -Situation ist nicht ganz klar einzuordnen. Einerseits gehört sie ebenso zum Reiz-Reaktions-Lerntypus wie die $k>2$ -Aufgaben, andererseits kann sie von der Vp auch als ein Miniaturproblem betrachtet werden: Bei f-Items sieht die Vp z.B. die Möglichkeit, auf die richtige Alternative zu "schließen".

Diese Überlegung wird durch Arbeiten ($k=2$) von z.B. SPENCE et al. (1963), SPENCE (1966) und LOVELACE (1966) gestützt. Die Arbeiten sind an sich klassische "blank-trial"-Untersuchungen, in denen das Problem der Bedeutungsauffassung von b durch die Vpn auftaucht (s. 2.3.1). Die Tendenz der Ergebnisse geht aber dahin, daß Symmetrie von r und f auftritt, falls nur die dargebotene Information adäquat ausgenutzt wird. Das wurde bei diesen Autoren dadurch erreicht, daß entweder den Vpn die Bedeutung von b ausdrücklich erklärt wurde oder daß Studenten als Vpn teilnahmen (s. auch die angenähert symmetrischen Ergebnisse aus Experiment I auf S. 106 und 108 für $k=2$).

Für das weitere wird jedenfalls in $k=2$ -Situationen eine vollständige, also symmetrische Informationsausnutzung der Rück-

wmeldungen r und f vorausgesetzt. - Die zusätzlich zu den oben angeführten Untersuchungen in BLÖSCHL (1969, S. 145) zusammengestellten Arbeiten zur relativen Effizienz in Alternativwahlsituationen unterscheiden sich in ihrer experimentellen Struktur zu sehr von den hier zu berichtenden Experimenten, um als weitere Ausgangspunkte für Hypothesen dienen zu können.

Für die Hauptargumentation in 6.3.3 bezüglich der Herstellung von symmetrischen und asymmetrischen Versuchssituationen entsprechend Fragestellungen (4) und (5) wird besonders die $k > 2$ -Aufgabe diskutiert werden, da sich hier ein interessanter Brückenschlag zur THORNDIKEschen Position andeutet.

Die $k=2$ -Aufgaben werden mehr die Funktion von Kontrollbedingungen übernehmen. In $k=2$ -Situationen verstärkt zu beachtende unterschiedliche Strategien der Vpn können nur am Rande mitdiskutiert werden (s. Experiment V in 7.2.2). Eine ausführlichere Behandlung dieser Strategien würde den Rahmen der vorliegenden Arbeit überschreiten (s. dazu HERRMANN u. STAFF, 1973).

Begriffsbildung wird aus zeitökonomischen Gründen nicht weiter behandelt. Da aus den Voruntersuchungen, Experimente I und II, und aus der Semesterarbeit von BERGHÄUSER u. SCHRÖDER (1972) schon Daten aus $k=2$ - und $k=4$ -Aufgaben vorliegen, sollen diese beiden Aufgabentypen im folgenden untersucht werden. Es werden demnach nur symmetrische und zugunsten von r asymmetrische Situationen diskutiert. Durch den Verzicht auf Begriffsbildungsprobleme geht allerdings leider die Möglichkeit der Herstellung einer asymmetrischen Situation zugunsten von f verloren.

6.2.3 "Nichts" als neutraler Bezugspunkt?

In diesem Kapitel soll nicht die in 2.3.1 angeschnittene Diskussion der Bedeutung von b im Kontrast zu r und f weiter fortgeführt werden. Es geht vielmehr um die numerische

Berechnung der relativen Effizienz von r und f bezogen auf bb , also das Wiederholungsniveau nach b .

Auf S. 44 war als ein aus den BUCHWALD-Profilen zu bestimmendes relatives Effizienzmaß: $re(b) = (rr-bb)/(rr-ff)$ vorgeschlagen worden, in Übereinstimmung mit BUCHWALDs Argumentationen (1962). Dieses Maß hat nun einige Nachteile:

1. Es ist abhängig von den Zufallsschwankungen der Größen (rr, bb, ff) , die sich wegen der Differenzen- und Quotientenbildung sehr empfindlich im numerischen Wert von $re(b)$ bemerkbar machen.

2. Nach BUCHWALD (1962) und HUMPHREYS (1972) deutet sich die Vermutung an, daß die Wirkung von b , die sich in bb bzw. in $re(b)$ ausdrückt, noch von k abhängt. Dadurch würden die Folgerungen BUCHWALDs (1962) aus seinen Profilen bezüglich der unterschiedlichen Effizienz von r und f in Abhängigkeit von k neu zu überdenken sein (s. dazu auf S. 25); es ist nicht zu vergessen, daß die Strategieneklärung zugunsten von f bei $k=2$ ad hoc gemacht wurde.

Spekulativ läßt sich die k -Abhängigkeit von $re(b)$ in dem Parameter g lokalisieren, wie durch eine quasi-quantitative Betrachtung mit Hilfe des BM illustriert werden soll; "quasi-quantitativ" deshalb, weil ja das BM nicht mehr als streng gültig vorausgesetzt werden kann.

Die Alternativenzahl k macht sich in rr , ff und bb durch das Zufallswiederholungsniveau $1/k$ bemerkbar (s. S. 98: Gleichungen (1a)' bis (1c)'). Durch die Differenzen- und Quotientenbildung bei der Berechnung von $re(b)$ fallen diese Zufallsanteile heraus, und es bleibt theoretisch übrig: $re(b) = 1-g$ bzw. $= 1-\hat{g}$.

Falls man annimmt, daß der Beharrungsparameter g sich additiv aus einem unsystematischen Zufallsanteil $1/k$ und einem systematischen "wahren" Beharrungsanteil g_w zusammensetzt, d.h. $g = 1/k + g_w$, ergibt sich: $re(b) = 1 - 1/k - g_w$, also $= re(b, k)$. Diese Aufspaltung in $1/k$ und g_w wird durch die Ergebnisse des Experiments I zur Schätzung des Parameters g nahegelegt (s.

S. 109: $\hat{g}_{k=2}=0,67$; $\hat{g}_{k=4}=0,47$).

Man betrachte nun $re(b,k)=1-1/k-g_w$ und erinnere sich (S. 45), daß $re(b) < 50\%$ bzw. $re(b) > 50\%$ Asymmetrie zugunsten von f bzw. von r bedeutete.

Man sieht sofort, daß bei $k=2$, also $1/k=50\%$, $re(b,k=2)=100\%-50\%-g_w < 50\%$ wird. (Es ist dabei $g_w > 0$ angenommen, was im Rahmen der bisherigen Überlegungen und Ergebnisse nicht unvernünftig erscheint.) Das bedeutet Asymmetrie zugunsten von f, wie es auch von BUCHWALD (1962) für $k=2$ gefunden wird. Der Formel für $re(b,k=2)$ nach ergibt sich unter einer $k=2$ -Bedingung überhaupt keine Chance für Asymmetrie zugunsten von r.

Zu vergleichen ist dagegen $re(b,k=4)=100\%-25\%-g_w=75\%-g_w$. Hier besteht durchaus die Möglichkeit für Werte von $re(b,k=4) > 50\%$, also Asymmetrie zugunsten von r, wie es auch gefunden wird.

Das Gesamtargument kann kurz und mit etwas anderen Worten so ausgedrückt werden: Die allgemeine Zufallsabhängigkeit in den Wiederholungsniveaus nach r, f und b schlägt bei bb über den Parameter g unter $k=2$ -Bedingungen mit 50% mehr durch als unter $k=4$ -Bedingungen mit nur 25%. Auf jeden Fall ist diese k-Abhängigkeit von $re(b)$ hinderlich für die Bearbeitung des Problems der relativen Effizienz von r und f.

3. Es sei an die einschränkende Bemerkung BUCHWALDs (siehe auf S. 29) erinnert und daran, daß das bb-Wiederholungsniveau den Zweck hat, den g-Mechanismus auszuschalten (S. 33f.). Dieser Umweg zur Bestimmung der relativen Effizienz von r und f über b ist also ein Produkt der experimentellen Vorgehensweise, nämlich Wiederholungsdaten zu verwenden. Ein direkterer Weg wäre wünschenswert.

4. Abschließend bleibt immer noch das Problem, inwieweit die Vpn die Einführung von b in der Instruktion verstehen.

Es wird in 6.3.4 ein Effizienzmaß vorgeschlagen werden, das nur von den Schwankungen *z w e i e r*, zu r und f gehörenden Größen abhängt. Die Einführung von b könnte dadurch überflüssig werden, und es würden die Nachteile 1. bis 4. vermieden.

6.3 Neuer Ansatz

6.3.1 Modifikation des BUCHWALD-Modells

Nachdem also das BM als striktes mathematisches Modell abgelehnt worden ist, sollen jetzt übersichtlich die schon an verschiedenen Stellen angedeuteten Modifikationen des BM zusammengestellt werden. Dieses modifizierte BM wird den Rahmen ergeben, in dem die folgenden Diskussionen verlaufen werden (in Kombination mit dem in 6.3.2 vorzustellenden Modell). In groben Zügen wurde das modifizierte BM schon in Kapitel 4 bei Ableitung der Latenzzeithypothesen umrissen, da ja beabsichtigt war, diese Hypothesen allgemeiner und unabhängiger von der strengen Gültigkeit des BM aufzustellen.

Die Abänderungen ergeben sich im wesentlichen dadurch, daß die Forderung nach striktem Nonrehearsal fallengelassen wird. Dadurch ist eine Einbeziehung der Informations- und Strategieneklärung möglich (S. 135). Es handelt sich also gewissermaßen um eine Kombination der Ansätze BUCHWALDs von 1962 und 1969.

B e i b e h a l t e n bleibt: die gedächtnispsychologische Interpretation der "Verstärkungs"-Wirkung von r und f . Über die (Un-)Abhängigkeit der Gedächtnisprozesse S-R und S-O ist damit noch nichts gesagt. - Weiterhin kann die Erklärung der Asymmetrie bezüglich $1/k$ zuungunsten von f durch den g -Mechanismus beibehalten werden. Nach Einführung des neuen Effizienzmaßes in 6.3.4 braucht auf diesen Mechanismus allerdings nicht mehr zurückgegriffen werden.

F a l l e n g e l a s s e n wird: die Unabhängigkeitsannahme $U3$ (S. 34). Weder BUCHWALDs Ergebnisse noch die der vorliegenden Arbeit geben irgendeinen direkten Hinweis auf die Gültigkeit dieser sehr starken Annahme.

M o d i f i z i e r t im Sinne BUCHWALDs (1962) werden: die

Unabhängigkeitsannahmen U1 und U2. Die Erinnerungen an R bzw. O brauchen nicht mehr unabhängig von der Art der Rückmeldung O zu sein. Der Einfachheit und Kürze halber in den ursprünglichen Parametern des BM ausgedrückt könnten die Verhältnisse ungefähr so aussehen: für $k > 2 \rightarrow c_r > c_f$ und $1-u_r > 1-u_f$; für $k=2 \rightarrow c_r < c_f$ und $1-u_r < 1-u_f$.

Die Beziehungen für $k > 2$ werden auch durch die Ergebnisse der Experimente I und II gestützt. - Die Beziehungen für $k=2$ wurden allerdings speziell als Folgerung aus BUCHWALDs ad-hoc-Hypothese bezüglich der asymmetrischen Konzentration zugunsten von f unter $k=2$ -Bedingungen hingeschrieben. Diese ad-hoc-Hypothese ergab sich nur bei Beziehung der relativen Effizienz auf bb. Die dazugehörige Problematik wurde schon in 6.2.3 (Punkt 2) erörtert.

Die Ergebnisse des Experiments I (R- und O-recall für $k=2$; s. S. 106 und 108) und die auf S. 137 zitierten Untersuchungen legen symmetrische Beziehungen näher, also im Sinne der reinen Informationshypothese ohne zusätzliche Strategieneklärung, d.h.: $c_r = c_f$ und $1-u_r = 1-u_f$ für $k=2$.

Daß bisher nichts zu c_b und $1-u_b$ gesagt wurde, ergibt sich aus der Diskussion in 6.2.3. Daraus folgt als letzte "Modifikation":

A n g e z w e i f e l t wird: die "neutrale" Bedeutung von b. Nichtsdestoweniger wird die Rückmeldung b in den weiteren Experimenten aus Vergleichbarkeitsgründen mitgeführt werden.

Kurz zusammengefaßt ist für das weitere wichtig: Die unterschiedlichen Rückmeldungen beeinflussen die Leistung, indem sie 1. während des Lernvorgangs (Einspeicherung) unterschiedliche Aufmerksamkeitszuwendungen bzw. unterschiedliches Rehearsal bewirken (BUCHWALD, 1962) u n d 2. im Prüfdurchgang (Erinnerung) die Entscheidung beeinflussen (BUCHWALD, 1969).

5.3.2 Ein ergänzendes Gedächtnismodell zur Verstärkung

Das im folgenden darzustellende Modell nimmt Bezug auf "Informationstheoretiker" wie TOLMAN (1932), ESTES (1969) und BUCHWALD (1969). Auch das modifizierte BM (s. 6.3.1) fügt sich in den relativ weiten Rahmen dieses Modells. Es handelt sich dabei um das Gedächtnismodell von ATKINSON u. SHIFFRIN (1968). Eine insbesondere auf das Verstärkungsproblem zugeschnittene Darstellung, der hier gefolgt werden soll, geben ATKINSON u. WICKENS (1971).

Zu Beginn wird, in Gegenüberstellung zu THORNDIKE, der Rahmen in Begriffen der Informationsverarbeitung gekennzeichnet.

"Verstärkung" (reinforcement) wird in aller Kürze definiert (S. 67): "... learning involves the transfer of information.. into some form of memory store that can hold it until it is needed later. Reinforcement is a modulation of this information flow."

Weitere kurze Kennzeichnungen sind (S. 68): "... a reinforcing event plays two... separate and distinct roles in determining the storage of information: an informational role and an attentional role... The attentional component of reinforcement... is closely related to conventional ideas of reward..." (s. dazu auch HARLEY, 1965a,b).

In diesem Modell werden Strukturkomponenten (structural features) und Kontrollprozesse (control processes) unterschieden.

Die Strukturkomponenten sind nicht zu beeinflussen, sie sind gewissermaßen "fest verdrahtet". Zu ihnen gehören das Sensorische Register, der Kurzzeitspeicher und der Langzeitspeicher, deren allgemeine Beschreibung sich wohl erübrigt. Die strikte Festlegung auf genau drei Speicher ist im übrigen für die weitere Diskussion nicht kritisch (s. dazu WICKELGREN, 1970; BERNBACH, 1970).

Die Kontrollprozesse sind dagegen sehr variabel. Sie hängen stark von der Einstellung der Vp ab, auch von ihrem Ermüdungszustand, usw. Ebenso werden sie von der Art der Lernaufgabe,

den Rückmeldungen und den Instruktionen in hohem Maße beeinflusst. Zu den Kontrollprozessen gehören Kodieren, Rehearsal, Transfer vom Kurzzeit- zum Langzeitspeicher, Suchstrategien usw.

Ein besonderer Kontrollprozeß ist die temporäre, von der jeweiligen Aufgabe abhängige Einrichtung eines "rehearsal buffer" im Kurzzeitspeicher. Dieser "rehearsal buffer" hat eine begrenzte Kapazität.

Er dient u.a. dazu, Information, die im Augenblick aktuell benötigt wird, für eben die notwendige Zeit zur Bewältigung der vorliegenden Aufgabe festzuhalten.

Ferner kann er dazu dienen, Information, die aufgrund eines anderen Kontrollprozesses als für die Zukunft relevant beurteilt worden ist, so zu kodieren und in den Langzeitspeicher zu transferieren, daß diese Information möglichst sicher gespeichert wird und effektiv wiedergefunden werden kann.

Eine weitere Charakteristik der Kontrollprozesse ist das sehr variable Ausmaß der Bewußtheit, mit der diese Prozesse durchgeführt werden, angefangen von höchst bewußter Kontrolle bis zum anderen Ende einer mehr automatischen Informationsverarbeitung. Solche automatisierten Kontrollprozesse rücken damit schon in die Nähe von "fest verdrahteten" Strukturkomponenten (s. dazu auch SHIFFRIN u. ATKINSON, 1969, S. 184 u. 186).

Die beiden genannten Aufgaben des "rehearsal buffer" und die Möglichkeit automatisierter Kontrollprozesse werden in 6.3.3 zur Argumentation angeführt werden.

Diese sehr knappe Darstellung des Gedächtnismodells von ATKINSON u. SHIFFRIN genügt für die Zwecke dieser Arbeit. Inwieweit geht es nun über das modifizierte BM hinaus, welchen zusätzlichen Nutzen bringt es?

Auch das modifizierte BM beinhaltet Kontrollprozesse: z.B. die bewußte, determinierte Entscheidung, bei Erinnerung an die Antwort R und die Rückmeldung f im Prüfdurchgang diese Antwort durch eine andere zu ersetzen, oder die vielleicht nicht ganz so voll bewußte, nicht ganz so konsequent durch-

geführte Strategie, sich im Lerndurchgang bei $k > 2$ -Aufgaben besonders auf die r-Items zu konzentrieren, d.h. also ein besonders intensives Rehearsal bei r-Items.

Diese Kontrollprozesse beziehen sich aber bei BUCHWALD nur auf den engen Bereich seiner Experimente und sind teilweise erst nachträglich postuliert worden.

Das augenblicklich behandelte Gedächtnismodell dagegen enthält als Grundzug von vornherein eine große Vielfalt von Kontrollprozessen, die aber durch den Rahmen der Strukturkomponenten so weit integriert werden, daß diese Vielfalt nicht in theoretische Bedeutungslosigkeit ausufert. Ferner werden in einer Reihe von Experimenten diese Kontrollprozesse und die von ihnen empfindlich abhängige Wirksamkeit der Rückmeldungen näher demonstriert bzw. empirisch untermauert.

Es ist das Ziel des folgenden Abschnitts, einen als sehr stark automatisiert angenommenen Kontrollprozeß zu beschreiben. Die Nützlichkeit dieses speziellen hypothetischen Kontrollprozesses bzw. die aus ihm abgeleiteten Hypothesen werden in den folgenden Experimenten (hauptsächlich Experiment III) einer Prüfung unterworfen werden.

6.3.3 Automatisierter Kontrollprozeß und Asymmetrie in Mehrfachwahlaufgaben

Die folgende Argumentation bezieht sich schwerpunktmäßig auf $k > 2$ -Situationen. Als Folgerungen werden sich Bedingungen ergeben, unter denen Asymmetrie bzw. Symmetrie von r und f zu erwarten ist (s. Fragestellungen (4) und (5) auf S. 5).

Der Anstoß zu folgenden Überlegungen kam von Experiment II bzw. von der dort (s. S. 125, Frage 2) berichteten auffälligen Schwierigkeit bei den Vpn, sich in einer $k=4$ -Situation trotz "symmetriebetonender" Instruktion gleichmäßig auf r- und f-Items während des Lerndurchgangs konzentrieren zu können. Dazu ein Zitat (REITMAN, 1970, S. 505): "If subjects find it impossible to carry out certain classes of tasks,

it may be reasonable to search for some basic limitation or constraint that is responsible."

Es sei zunächst noch einmal daran erinnert, wie die Wirkung von Bestrafung sich nach THORNDIKE darstellt (s. dazu das auf S. 14 oben angeführte Zitat): Sie ruft andere Reaktionen hervor, unter denen unter Umständen dann auch die richtige ist.

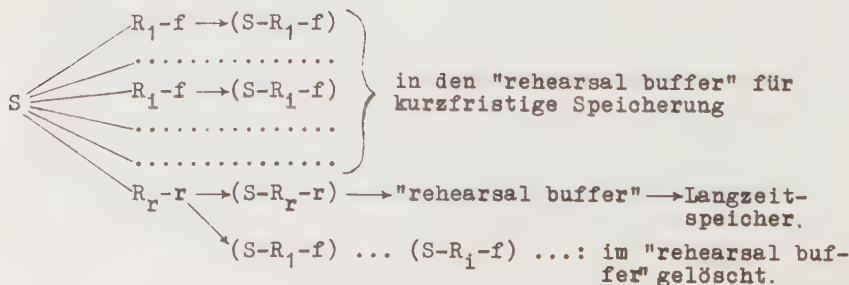
Die Entstehung eines automatisierten Kontrollprozesses, der in $k > 2$ -Situationen einsetzt, sei nun versucht, plausibel zu machen.

Im Laufe der allgemeinen Lerngeschichte von menschlichen Individuen werden viele Trial-and-error-Situationen, also speziell $k > 2$ -Probleme, bewältigt. Dabei kann, besonders bei motorischem Lernen, k sehr groß sein. (Die Diskussion um die relative Effizienz von r und f braucht durchaus nicht auf verbales Lernen beschränkt zu werden; s. S. 13.)

Als Illustration kann sich für die folgende Argumentation z.B. diese $k > 2$ -Situation vorgestellt werden: Die Reizsituation S ist das Armaturenbrett eines fremden Autos, die Scheibenwischer stehen still, und es regnet. Man möchte "richtig" reagieren, nämlich den Knopf betätigen (R_r), der zum Erfolg (r) führt, d.h. zur Bewegung der Scheibenwischer.

Das Drücken, Ziehen oder Drehen von Knöpfen sind gut bekannte Reaktionen, es braucht jetzt also nur unter mehreren Knopfbetätigungsreaktionen (R_1) die richtige (R_r) gefunden zu werden.

In der Darstellungsweise des oben (s. 6.3.2) beschriebenen Gedächtnismodells könnte die Bewältigung der $k > 2$ -Situation dann so aussehen:



Die Reaktion R_1 könnte z.B. das Betätigen des Knopfes für die Scheibenwaschanlage sein, also f. R_1 wird dann kurzfristig als f-Reaktion im "rehearsal buffer" gespeichert, damit sie in der augenblicklichen Problemsituation nicht nochmals wiederholt wird. Ebenso geschieht es mit den anderen f-Reaktionen R_1 , bis durch Zufall die richtige Reaktion R_r auftritt. Dieser wendet sich nun die volle Aufmerksamkeit zu, d.h. $S-R_r-r$ wird im "rehearsal buffer" geeignet kodiert und in den Langzeitspeicher transferiert. Zuvor werden sämtliche andere Sequenzen ($S-R_1-f$) im "rehearsal buffer" gelöscht, weil sie ja im Augenblick nicht mehr gebraucht werden.

Es wurde hier nur der "reine" Fall beschrieben. Natürlich kommt es vor, daß eine f-Reaktion in derselben Situation noch ein zweites Mal wiederholt wird oder daß z.B. R_1 als richtige Reaktion für die Betätigung der Scheibenwaschanlage in den Langzeitspeicher eintritt.

Bei alltäglichen $k > 2$ -Problemen braucht man sich also im allgemeinen falsche Reaktionen nur kurzfristig zu merken. Es wäre unökonomisch, sich für solche Situationen die f-Reaktionen über längere Zeit einprägen zu wollen. - Die (Nicht-)Berührung eines heißen Ofens stellt kein $k > 2$ -, sondern ein $k=2$ -Problem dar.

$k > 2$ -Problemsituationen sind nun, von Geburt an, in der allgemeinen menschlichen Lerngeschichte sehr häufig: Das menschliche Individuum hat g e l e r n t, wie es in solchen Situationen z u l e r n e n hat, nämlich, sich f-Reaktionen nur kurzfristig einzuprägen. Es ist ein relativ stark automatisierter globaler Kontrollprozeß entstanden. "Global" bezieht sich dabei auf die Bewältigungsweise der gesamten $k > 2$ -Situation.

Sobald also ein Individuum auf eine $k > 2$ -Situation trifft, wird dieser Kontrollprozeß angestoßen, der bewirkt, daß f-Reaktionen ohne allzu große Aufmerksamkeitszuwendung kurzfristig abgespeichert werden. Da der "rehearsal buffer" nur eine begrenzte Kapazität hat, können f-Reaktionen zugunsten von r-Reaktionen, die zu anderen Reizsituationen gehören, leicht

wieder verdrängt (gelöscht) werden.

Die Folgerung ist nun:

Gegeben sei eine $k > 2$ -, z.B. $k=4$ - Situation (genauer: eine Reihe von $k > 2$ -Situationen, wie z.B. bei BUCHWALD oder in den Experimenten I und II), in der per Instruktion verlangt wird, sich bei allen Reizwörtern gleichmäßig auf r- und f-Reaktionen zu konzentrieren, diese also über eine längere Zeit im Gedächtnis bis zum Prüfdurchgang zu behalten.

Dann besagt die hier im weiteren zu vertretende Hypothese, $H(\text{akt-4})$ ¹⁾, daß der über viele Jahre in zahlreichen natürlichen Lernsituationen eingeschliffene Kontrollprozeß sich gegenüber der relativ künstlichen Instruktion (nämlich sich f-Reaktionen für längere Zeit zu merken) durchsetzen wird. Empirisch wird sich also trotz scharfer symmetrie betonender Instruktion und noch zusätzlicher symmetrieunterstützender experimenteller Manipulation ("Verbalisieren"; s. 6.3.4) eine Asymmetrie zugunsten von r ergeben.

Der wesentliche Unterschied zwischen den eben beschriebenen Alltags- bzw. Experimentalsituationen, durch den obige Hypothese überhaupt möglich wird, ist der, daß in Alltagssituationen unter der "Entdeckungs-Methode" (s. S. 15) gelernt wird, die THORNDIKE- und BUCHWALD-Experimente aber "vanishing situations" darstellen, in denen unter der "Nichtkorrektur-Methode" gelernt wird.

Zu der Künstlichkeit der letzteren Situationen schreibt POSTMAN (1962, S. 360), nachdem er vorher ihren Wert gewürdigt hat: "...such situations are unusual. The results obtained in 'vanishing situations' cannot be generalized to the effects of punishment in general...".

Bevor nun die oben aufgestellte Hypothese $H(\text{akt-4})$ weiter präzisiert wird, seien nun noch einige Anmerkungen hinzugefügt.

1) Zur abkürzenden Schreibweise: "akt" bedeutet "aktives" Lernen im Sinne von "performing" im Gegensatz zu "beobachtendem" Lernen ("observing"); "4" bedeutet $k=4$.

Der beschriebene relativ stark automatisierte Kontrollprozeß hat gewisse Ähnlichkeiten mit Phänomenen, bei denen man Begriffe anwenden würde wie z.B. Gewohnheit (habit), Lernhaltung bzw. Erlernen des Lernens (learning set, learning to learn), Verstärkungsvorgeschichte (reinforcement history). Versuchsweise könnte man eine gewisse Reihenfolge aufstellen, wobei der ordnende Gesichtspunkt die Zeitdauer ist, die für die Etablierung der "Lernhaltungen" benötigt wird. Die folgenden vier Beispiele sind nach Anwachsen dieser Zeitdauer geordnet:

1. BUCHWALD (1959a,b; "acquisition-of-reinforcement-value-hypothesis" bezüglich der Bedeutung von b in Kontrastierung zu r und f); 2. HARLOW (1949; learning set, learning to learn); 3. BARON (1966; social reinforcement history); 4. Automatisierter Kontrollprozeß, wie oben beschrieben. - Es sei nochmals betont, daß diese Randbetrachtung nur der Versuch einer Einordnung ist.

Welchen theoretischen Standort hat nun dieser automatisierte Kontrollprozeß? Einerseits wurde er im Rahmen eines kognitiv orientierten Gedächtnismodells in Begriffen der Informationsverarbeitung dargestellt¹⁾, andererseits wird durch die Verwendung des Wortes "automatisiert" stark an THORNDIKES "direkte, automatische und unvermeidbare" Wirkung von Verstärkungen, speziell von Belohnungen auf r-Reaktionen, erinnert (s. POSTMAN, 1962, S. 396).

Die wesentlichen Unterschiede zum THORNDIKESchen Standpunkt bestehen darin, daß beim automatisierten Kontrollprozeß kein Bezug auf Emotionen genommen wird und daß seine Wirkung sich hauptsächlich auf die automatische Vernachlässigung der f-Reaktionen bezieht.

Nichtsdestoweniger wird hier ein gewisser Brückenschlag zur THORNDIKESchen Position gesehen, zumal die Vorhersage,

1) Zur prinzipiellen Möglichkeit automatischer Wirkungen von Verstärkungen innerhalb eines kognitiven Systems s. auch NUTTIN u. GREENWALD (1968, S. 150).

Asymmetrie zugunsten von r , von beiden Standpunkten aus dieselbe ist ($k > 2$ vorausgesetzt). Es wird sich allerdings der allgemeine Rahmen der Informationsverarbeitung unter bestimmten Umständen ($k=2$) als der umfassendere erweisen.

In bezug auf $k=2$ -Situationen ist der automatisierte Kontrollprozeß entsprechend seiner Einführung prinzipiell nicht anwendbar. Es spricht aber nichts dagegen, in diesem Fall die rein kognitive Informationshypothese den Vorhersagen zugrunde zu legen. Danach wird gleiche Informationsübermittlung durch r und f angenommen und also Symmetrie von r und f vorausgesetzt.

Nach THORNDIKE dagegen müßte sich auch hier Asymmetrie zugunsten r ergeben. Daß THORNDIKE sein "allgemeines" asymmetrisches Effekt-Gesetz aus $k > 2$ -Situationen abgeleitet hat, ist eben auch die Schwäche dieses Gesetzes.

Die Ergebnisse der $k=2$ -Gruppen aus Experiment III (s. 7.1) zeigen eindeutige Symmetrie an. Sie stellen eine Kontrolle für die Kognitivität der experimentellen Situation dar; damit können jedenfalls Erklärungen durch emotionale Wirkung von r und f nach THORNDIKE auch für die (asymmetrischen) $k=4$ -Gruppen ausgeschlossen werden.

Mit diesen Bemerkungen zur $k=2$ -Situation wird von der allgemeinen Diskussion wieder übergeleitet zur speziellen Anwendung auf das Problem der Herstellung (a-)symmetrischer Versuchssituationen. Eine Möglichkeit dazu ist schon bekannt, eben die Gegenüberstellung von $k=2$ - und $k > 2$ -Aufgaben. Eine zweite Möglichkeit soll jetzt vorgestellt werden, auf die schon in 6.1 hingewiesen wurde. - Es sei dazu noch einmal an die betont kognitiv gehaltene Ausrichtung der vorliegenden Untersuchung erinnert; es wird im folgenden also nicht auf Lernsituationen des Alltags Bezug genommen, die mit starken Emotionen verbunden sind.

Die Begründung des automatisierten Kontrollprozesses hängt in hohem Maße von der großen Zahl der $k > 2$ -Probleme ab, die von den menschlichen Individuen im Laufe des Lebens bewältigt werden, und zwar durch "aktives" Ausführen der (rich-

tigen und falschen) Reaktionen.

Sehr viel seltener sind $k > 2$ -Problemsituationen, in denen ein anderer (ein "Modell") bei ihrer Bewältigung "beobachtet" wird (s. dazu die frühere Besprechung von "aktivem" und "beobachtendem" Lernen in 5.1.1).

Beim "beobachtendem" Lernen häufig dagegen sind Situationen, in denen ein "Modell" (z.B. Eltern) vom Lernenden beim Ausführen der *s c h o n r i c h t i g e n* Reaktion beobachtet wird. Diese Situationen kann man daher nicht mit der Nichtkorrektur-Methode in Verbindung bringen. Es wird hypothetisch angenommen, daß unter "beobachtenden" Lernbedingungen sich kein oder ein nur schwach automatisierter Kontrollprozeß entwickelt hat.

Eine weitere *H y p o t h e s e*, $H(\text{beob-4})^1$, ist also, daß in $k > 2$ -($k=4$ -)Situationen, z.B. in den augenblicklich betrachteten Experimenten, in denen "beobachtend" gelernt werden soll, der oben beschriebene Kontrollprozeß überhaupt nicht oder nur schwach einsetzt, sich jedenfalls gegen eine scharfe symmetriebetonende Instruktion nicht durchsetzen kann. Trotz $k=4$ wird Symmetrie in der Wirkung von r und f vorausgesagt.

Zwei Hauptvorteile dieser vermuteten zweiten Möglichkeit, Symmetrie herzustellen, sind:

1. Wird in einer $k=4$ -Situation eine symmetriebetonende Instruktion gegeben und ergibt sich dann unter Beobachtungsbedingungen Symmetrie, unter aktiven Bedingungen Asymmetrie zugunsten r , so kann der letztere Effekt nicht mehr durch ein Mißverständnis der Instruktion durch die V_{pn} erklärt werden (s. dritte Erklärungsmöglichkeit in 5.3.4, S. 127).
2. Entsprechend Fragestellung (3) auf S. 3 sind Latenzzeitmuster aus symmetrischen bzw. asymmetrischen Situationen miteinander zu vergleichen. Erzeugt man den Übergang von Asymmetrie zu Symmetrie durch Wechsel von $k > 2$ - zu $k=2$ -Aufgaben (unter "aktiven" Bedingungen), so kann unter Umständen

1) S. Fußnote auf S. 148.

eine starke, wenn auch für vorliegende Zwecke irrelevante k-Abhängigkeit der Latenzzeiten die Interpretation stören. Diese Schwierigkeit entfällt, falls man Symmetrie bzw. Asymmetrie durch Gegenüberstellung von "beobachtendem" und "aktivem" Lernen in n e r h a l b der k=4-Bedingung erzeugt.

Zur weiteren Begründung und Absicherung der oben aufgestellten Hypothesen bezüglich des wesentlichen Unterschiedes zwischen aktivem und beobachtendem Lernen in $k > 2$ -Situationen sollen noch einige Bemerkungen angefügt werden, die weiter oben den Fluß der Darstellung gestört hätten.

Soweit die früher erwähnten Arbeiten zum Beobachtungslernen (s. 5.1.1; insbesondere HILLIX u. MARX, 1960; MARX u. WITTER, 1972) vergleichbare Probleme behandeln, betonen sie hauptsächlich die relativ größere Fähigkeit der "Beobachter" (observer), die Informationen, die ihnen vom V1 geboten werden, zu verarbeiten, ungestörter z.B. durch eventuell nach f auftretende Emotionen. Dieses entspricht der Voraussage der Hypothese H(beob-4): Symmetrie bei Beobachtung und k=4, insofern als sich die symmetriebetonende Instruktion bei "Beobachtern" besser durchsetzen kann als bei "aktiven" Lernern. - Diese Anmerkung ist allerdings nur dann von Interesse, falls überhaupt unter Beobachtungsbedingungen ein wie oben beschriebener automatisierter Kontrollprozeß einsetzen sollte.

Über das bisher Gesagte, einschließlich der zitierten Literatur, hinaus könnte es sein, daß beim Übergang von "aktivem" zu "beobachtendem" Lernen bei k=4-Aufgaben, also nach obigen Hypothesen (akt-4 und beob-4) von Asymmetrie zu Symmetrie, Effekte auftreten, die außerhalb der augenblicklichen Betrachtungsweise liegen. Solche Effekte, z.B. komplizierte Wechselwirkungen, könnten die bisherigen Überlegungen sehr in Frage stellen.

Es scheint noch eine experimentelle Kontrolle angebracht, bei der es auch den Übergang von "aktivem" zu "beobachtendem" Lernen gibt, nicht aber gleichzeitig auch den vermuteten Wechsel von Asymmetrie zu Symmetrie. In diesem Fall wäre es möglich, den "aktiv $\rightarrow$ beobachtend"-Vorgang ungestört vom

(A-)Symmetrie-Problem zu analysieren. Diese Kontrollfunktion kann durch eine "aktive $k=2$ "- und eine "beobachtende $k=2$ "-Gruppe übernommen werden. Aufgrund der rein kognitiven Informationshypothese wird hier auf jeden Fall Symmetrie vorausgesagt.

Abschließend noch eine Klarstellung: Es war auf S. 145 gesagt worden, daß die Ergebnisse des Experiments II den Anstoß zu den Überlegungen gegeben haben, die zur hypothetischen Aufstellung eines automatisierten Kontrollprozesses und zu den Hypothesen H(akt-4) und H(beob-4) führten. Die Hypothese H(beob-4) sagt nun Symmetrie bei $k=4$ -Aufgaben unter Beobachtungsbedingungen voraus. In Experiment II hatte sich aber gerade unter diesen Bedingungen Asymmetrie ergeben.

Die versuchsweise Auflösung dieses anscheinenden Widerspruchs besteht im Hinweis auf die vier (bzw. fünf) Lerndurchgänge in Experiment II, durch die in zwar nicht ganz geklärter Weise eine Asymmetrie zugunsten r zustandegekommen sein kann (s. S. 127: zweite Erklärungsmöglichkeit).

Das Experiment II, insbesondere das postexperimentelle Interview, war also nur der Anlaß zu obigen Überlegungen, es war nicht der Ausgangspunkt, von dem eine strenge Ableitung zu den Hypothesen H(akt-4) und H(beob-4) führte.

Da im Hauptexperiment (III; nur e i n Lerndurchgang) die Hypothese H(beob-4) starke Unterstützung findet, spricht rückblickend viel für die zweite Erklärungsmöglichkeit bezüglich Experiment II.

6.3.4 "Verbalisieren" im Prüfdurchgang

Während die bisherigen Abschnitte 6.3.1 bis 6.3.3 relativ theoretisch ausgerichtet waren, soll in diesem Abschnitt hauptsächlich eine neue Technik bezüglich der Durchführung des Prüfdurchgangs beschrieben werden. Daß unterschiedliche Techniken aber auch zu unterschiedlichen theoretischen Aussagen führen können, ist bekannt und wird für den vorliegen-

den Fall am Ende des Kapitels erörtert.

Es gab im wesentlichen zwei Anlässe für die neue Befragungsweise im Prüfdurchgang: 1. Das Bestreben, die im BM bzw. in seiner modifizierten Form postulierten Erinnerungs- und Entscheidungsprozesse unmittelbar sichtbar zu machen, und damit zusammenhängend 2. das Bedürfnis, die Komponenten in den Latenzzeitverteilungen separieren zu können.

Das BM (ebenso auch das Modell von HUMPHREYS, 1972) erklärt die beobachtbaren Daten, z.B. Wiederholungshäufigkeiten, durch eine Verknüpfung von hypothetischen Erinnerungsprozessen (Axiome A1 und A2) mit hypothetischen Entscheidungsprozessen (A3). Dieses Vorgehen ist durchaus regelgerecht.

Bei der Nachprüfung der Axiome ergaben sich aber, wie beschrieben, gewisse Schwierigkeiten. Diese sollten nun, zwar etwas weniger elegant, dafür aber sicherer gelöst werden, indem nicht nur das Ergebnis aus der Verknüpfung der Axiome beobachtet wurde, also die Wiederholungsdaten, sondern die Erinnerungs- und Entscheidungsprozesse selbst sichtbar gemacht wurden.

Die Befragungstechnik ist einfach.

Im Lerndurchgang habe die Vp auf das Reizwort TEIG mit 5 geantwortet und vom V1 als Rückmeldung f erfahren; die aufgetretene Sequenz ist also TEIG 5 f . Im Prüfdurchgang soll die Vp im Sinne der üblichen Instruktion, also entsprechend R-corr, antworten, anschließend aber "begründen", warum sie gerade so geantwortet hat, d.h. sie soll versuchen, sich an ihre Antwort R und die Rückmeldung O aus dem Lerndurchgang zu erinnern (entsprechend R- u n d O-recall). Im obigen Beispiel würde die Vp (bei korrekter Erinnerung) im Prüfdurchgang auf TEIG hin z.B. antworten: 3— wegen Erinnerung an 5 und f, oder kürzer: 3—5,f. (Bei k=4 könnte natürlich auch 7 oder 9 statt 3 gesagt werden.)

Ein anderes Beispiel: Im Lerndurchgang sei die Folge aufgetreten WEIN 9 b . Im Prüfdurchgang könnte dann die Vp so auf WEIN reagieren: 9—9,b. Die Vp hat also nach b ihre

Antwort wiederholt. Sie hätte auch anders reagieren können: 3—9, b. Man sieht die Möglichkeit, unmittelbar die allgemeine Reaktion der Vpn auf b festzustellen, d.h. den Parameter g , genauer g' , direkt zu schätzen (s. S. 30).

Im ganzen handelt es sich also um eine Kombination von R-corr-, R- und O-recall-Prüfdurchgängen, nach der Einteilung von BLÖSCHL (1969, S. 39; s. besonders S. 126-141) um eine gemeinsame Wiederholungs- und Behaltensprüfung. Auch bei POSTMAN u. ADAMS (1954, 1955) findet sich eine ähnliche Kombination. Dort wird aber erst ein Prüfdurchgang unter R-corr durchgeführt, dann folgt anschließend eine erneute Darbietung der Prüfliste unter R- und O-recall. In vorliegender Arbeit wird dagegen innerhalb eines Prüfdurchgangs jedes Item unmittelbar nacheinander unter R-corr, R- und O-recall abgeprüft.

Dieses Verfahren hat die Folge, daß die Entscheidungsprozesse der Vp, die zwischen den Erinnerungen an R und O und den (Nicht-)Wiederholungsantworten liegen, nicht verlorengehen. Daraus folgt als ein für das Thema dieser Arbeit wichtiger Vorteil, daß eine sehr enge Verknüpfung hergestellt wird zwischen den Latenzzeiten, die weiterhin für die (Nicht-)Wiederholungsantworten gemessen werden, und den Erinnerungs- und Entscheidungsprozessen, die zu eben diesen Antworten geführt haben.

Die bisherige Sprechweise (s. Kapitel 4) bezüglich der Antworten, die über verschiedene "Wege" der B-Diagramme zustandekämen, und der zugehörigen Latenzzeiten bleibt nicht mehr leer, sondern kann nun für jedes einzelne Item belegt werden.

Wie schon in Experiment I unter der R- bzw. O-recall-Bedingung ein Test des BM hinsichtlich der Unabhängigkeitsannahmen U1 bzw. U2 möglich war, so ist diese Gelegenheit auch jetzt gegeben. Darüber hinaus ergibt sich durch die gemeinsame Befragung nach R und O bei jedem einzelnen Item die Möglichkeit, auch U3 zu testen (s. S. 34) bzw. ein Maß für die Abhängigkeit von R und O zu bestimmen.

Anhand der oben angeführten Beispiele (Antworten der Vp im Prüfdurchgang auf TEIG mit 3—5,f bzw. auf WEIN mit 9—9,b) soll eine kurze Bemerkung zur Ausdrucksweise gemacht werden. Wie bisher werden die jeweils ersten Ziffern, also hier 3 und 9, als Wiederholungsdaten bezeichnet werden, die darauf folgenden Paare, also hier 5,f und 9,b, als R- und O-Erinnerungsdaten. Zusammengenommen, also 3—5,f und 9—9,b, soll von Verbalisierungsdaten gesprochen werden.

Naheliegender wäre auch gewesen, die Paare 5,f bzw. 9,b als Verbalisierungen zu den Antworten 3 bzw. 9 zu bezeichnen. Aber erstens sind die Bedeutungen der wie oben festgelegten Wiederholungsdaten (repetition measures) und (R-)Erinnerungsdaten (recall measures) schon durch die theoretische Kontroverse zwischen GREENWALD (1966) und POSTMAN (1966) so stark belastet und festgelegt, daß von dieser Bezeichnungsweise nicht abgewichen werden sollte.

Zweitens wäre für das obige Beispiel 9—9,b es nicht ausreichend, nur 9,b als Verbalisierung zu bezeichnen, da aus 9,b noch nicht eindeutig der Entschluß der Vp folgen würde, bei Erinnerung an 9 und b 9 zu wiederholen (die Vp hätte ja auch 3 sagen können).

Als nächstes wird eine Erweiterung des "Verbalisierens" im Prüfdurchgang beschrieben: Die Vp sollte bei den E r i n n e r u n g s antworten noch zusätzlich die subjektive Sicherheit (confidence) ihrer Erinnerungen angeben, und zwar g e t r e n n t für R und O. Es waren drei Konfidenzstufen vorgegeben: "sicher", "unsicher" und "vergessen". Falls z.B. im Lerndurchgang die Sequenz TEIG 5 f aufgetreten war, könnte die Vp im Prüfdurchgang auf TEIG z.B. so reagieren: "3—5 sicher, f sicher" oder "5—vergessen, f unsicher" oder "3—3 unsicher, r sicher" (die fälschliche Erinnerung an r kann von der Vp durchaus subjektiv als korrekt und sicher beurteilt werden!) oder "5—5 sicher, vergessen" usw.

Die zuletzt angegebene Verbalisierung "5—5 sicher, vergessen" demonstriert, warum die Kategorie "vergessen" eingeführt wurde. Es besteht hier die Möglichkeit, einen der Hauptargu-

mentationspunkte des BM, nämlich den g- (genauer g"-)Mechanismus, näher zu analysieren.

Durch die Kategorie "vergessen" ist die Identifikation der Situation, daß die Vp sich zwar an R, nicht aber mehr an O erinnert, besser gewährleistet als durch eine Instruktion, auf jeden Fall eine Erinnerung (r,f oder b), notfalls mit niedriger Konfidenzstufe anzugeben.

Durch den Vergleich von Antworten der Vp wie z.B. "3—3 sicher, b sicher" und "5—5 sicher, vergessen" läßt sich also unmittelbar die "starke Annahme" BUCHWALDS (s. S. 30f.) bezüglich $g'=g=g$ überprüfen.

Die Einführung der Kategorien "sicher" bzw. "unsicher" war dadurch begründet, daß die Ableitung der Latenzzeithypothesen in Kapitel 4 (s. besonders S. 70) sich hauptsächlich auf die "sicheren" (c,1-u)-Wege stützte. Von den Antwortzeiten über die "unsicheren" (c,1-u)-Wege war angenommen worden, daß sie sich nicht sehr von den langen Zeiten der GK unterscheiden würden. - Die Antworten über die "sicheren" (c,1-u)-Wege bzw. die zugehörigen Latenzzeiten können also prinzipiell identifiziert und daher die Begründung der Hypothesen direkt veranschaulicht werden.

Nach diesen vorbereitenden Bemerkungen soll nun ein neues Maß für die relative Effizienz von r und f eingeführt werden.

Im Lerndurchgang seien z.B. die Sequenzen SIEB 3 r und TEIG 7 f aufgetreten. Wenn dann die Vp im Prüfdurchgang auf SIEB mit "3—3 sicher, r sicher" und auf TEIG mit "5—7 sicher, f sicher" antwortet, dann sollen per definitionem, im gedächtnispsychologischen Rahmen des BM und der vorliegenden Arbeit, die Rückmeldungen r und f wirksam geworden sein.

Diese Definition bezieht sich also ausschließlich auf die "problemrelevanten" (c,1-u)-Wege in den B-Diagrammen (als problemirrelevant, zumindest aber als problemferner in Hinblick auf die relative Effizienz von r und f sind die (1-c)/k- bzw. c·u·g-Wege aufgrund ihres Zufalls- bzw. "bias"-

Charakters zu beurteilen).

Selbstverständlich könnte eine entsprechende Definition z.B. auch Fälle wie "3—3 unsicher, r unsicher" oder "5—7 sicher, f unsicher" umfassen. Es soll aber bei der vorgeschlagenen Definition bleiben, um die Einheitlichkeit der Argumentationen mit dem Latenzzeitbereich zu wahren. Im Übrigen zeigte sich später bei der Berechnung der a-posteriori-Wahrscheinlichkeiten ($p(\text{korrekt}/\text{Konfidenzstufe})$; s. z.B. MURDOCK, 1970, S. 293), daß die Kategorie "sicher" sich deutlich von der Kategorie "unsicher" abhob, die kaum das Zufallsniveau überschritt (s. 7.1.3).

Im folgenden sei "sicher" durch "!" abgekürzt und das Umschalten von einer falschen Reaktion R auf eine andere durch $\bar{R}$ gekennzeichnet. Falls r bzw. f wirksam geworden sind, können diese Fälle dann allgemein so geschrieben werden: "R-R!r!" bzw. " $\bar{R}$ -R!f!".

Eine Rückmeldung, r bzw. f, soll um so wirksamer heißen, je häufiger die obigen Fälle auftreten.

Die absolute Häufigkeit sei durch $n(\dots)$ abgekürzt, also $n(R-R!r!)$ und $n(\bar{R}-R!f!)$. Wird von der gleichen Anzahl r- bzw. f-Items (z.B. 12:12) ausgegangen, so bedeutet dann z.B. $n(R-R!r!) > n(\bar{R}-R!f!)$, daß r wirksamer als f ist.

In diesem Sinne wird als Maß für die relative Effizienz von r bezüglich f vorgeschlagen:

$$re(r!) = n(R-R!r!) / (n(R-R!r!) + n(\bar{R}-R!f!))$$

Enthält eine Lernliste ungleiche Anteile von r- und f-Items, müßten die absoluten Häufigkeiten $n(\dots)$ noch vorher durch die jeweilige Anzahl der r- bzw. f-Items dividiert werden. - Die relative Effizienz von f bezüglich r ist natürlich $re(f!) = 1 - re(r!)$.

Das neue Effizienzmaß $re(r!)$ ist sinngemäß ähnlich gebildet wie $re(b)$ auf S. 44. Es ist aber unabhängig von b, und es entfallen daher sämtliche in 6.2.3 aufgezählten Nachteile von $re(b)$ (s. dort insbesondere Punkt 3, S. 140, auf

den am Ende dieses Abschnitts noch einmal hinzuweisen ist). Numerisch ist $re(r!)$ mit $re(b)$ insofern vergleichbar, als auch bei $re(r!)$ gilt: $re(r!) = 50\%$ bei Symmetrie von r und f ; $re(r!) < 50\%$ bei Asymmetrie zugunsten von f ; $re(r!) > 50\%$ bei Asymmetrie zugunsten von r (s. S. 45).

Ferner wird auch der auf S. 128 diskutierte prinzipiell mögliche "ceiling"-Effekt vermieden.

Auf zwei mögliche Schwächen des neuen Effizienzmaßes ist hinzuweisen:

1. Die Bestimmung von $re(r!)$ über die Verbalisierungstechnik ist aufwendiger als die von $re(b)$, bei der nur Wiederholungsdaten erhoben werden müssen. Bei späteren Anwendungen aber, vorausgesetzt die Einführung von $re(r!)$ bewährt sich überhaupt in dieser Arbeit, können einige Vereinfachungen vorgenommen werden, die sowohl Zeitersparnis bringen als auch für die Vpn angenehmer sind: Die Einführung von b mit allen ihren Komplikationen kann unterbleiben und, falls nicht wieder Latenzzeitanalysen vorgenommen werden sollen, ist auch das Abgeben von Konfidenzurteilen nicht unbedingt nötig.

2. Bei der Definition von $re(r!)$ ist die gedächtnispsychologische Interpretation des "Verstärkungsvorgangs" nach BUCHWALD vorausgesetzt, und durch die angewandte Verbalisierungstechnik wird die Gültigkeit dieser Interpretation auch praktisch erzwungen (s.dazu S. 2). Dadurch ist der Nutzen der Einführung von $re(r!)$ zunächst auf die noch folgenden Experimente der vorliegenden Arbeit beschränkt. Das Ausmaß der speziellen Effekte, die möglicherweise durch das "Verbalisieren" erzeugt werden, soll in Experiment III durch zwei Kontrollgruppen ($k=2,4$) abgeschätzt werden, die ohne Verbalisierungsanweisung "konventionell" geprüft werden.

Dieser letztere Punkt ist wegen seiner Bedeutung noch weiter zu diskutieren. Das Problem ist also: Werden Vpn durch die Verbalisierungsinstruktion schon in den Lerndurchgängen so manipuliert, daß ihre Konzentration, ihre Strategien, ihre Informationseinspeicherung sehr stark abweichen von ihrem

sonstigen Lernverhalten unter "freien", also Nichtverbalisierungs-Bedingungen? Die Folgen einer solchen Manipulation könnten sein:

a) Die Gruppen "mit Verbalisieren" unterscheiden sich von jenen "ohne Verbalisieren" hinsichtlich des erzielten Lernniveaus. Die Richtung des Unterschieds kann zunächst nicht vorausgesagt werden.

b) Auch bei gleichem Lernniveau unterscheiden sich die tatsächlich stattgefundenen Einspeicherungs- und Suchprozesse erheblich zwischen den beiden Gruppen.

Während sich Punkt a) über die Steilheit der konventionellen rr - bb - ff- Profile kontrollieren läßt, würden Unterschiede nach Punkt b) sich in den Latenzzeitmustern bemerkbar machen. In Anbetracht des Hauptthemas dieser Arbeit ist die mögliche Verzerrung der Latenzzeitmuster eine kritische Stelle, an der die Einführung des "Verbalisierens" scheitern könnte.

Dazu folgende Überlegung: Sicherlich kann nicht angenommen werden, daß durch die Verbalisierungsinstruktion sich überhaupt nichts bei den Gedächtnis- und Entscheidungsprozessen ändert. Falls nun aber nicht gerade wesentlich verfälschende Effekte ins Spiel kommen, ist vielmehr zu vermuten, daß im Prüfdurchgang die endgültigen Antworten sorgfältiger überlegt werden, die "verdeckten" (covert) Gedächtnis- und Entscheidungsprozesse mehrere Male durchlaufen werden. Nach dem, was auf S. 61 zu der speziellen Effektvorhersage des Modells von LaBERGE (1962) und auf S. 125f. zum Unterschied der Latenzzeitmuster unter "speed"- und "nonspeed"-Anweisungen gesagt wurde, ist zu erwarten, daß sich das Latenzzeitmuster bei den Verbalisierungs-Gruppen zu längeren Zeiten verschieben und sich dabei auffächern wird. Unter "auffächern" ist dabei zu verstehen, daß sich die Abstände zwischen den Medianen vergrößern, ohne daß die Gesamtrangreihenfolge, bzw. schwächer, die lokalen Rangfolgen zwischen je zwei Medianen bis auf Zufallsschwankungen gestört werden.

Man könnte diesen Effekt, falls er auftreten sollte, kurz "Zeitlupeneffekt" nennen.

Dieser Zeitlupeneffekt kann, wenn überhaupt, in abgestufter Strenge auftreten: 1. Verschiebung und Auffächerung entsprechend einer linearen Transformation; 2. Erhaltung der Gesamt-reihenfolge der sechs Mediane; 3. lokale Erhaltung bestimmter Beziehungen zwischen je zwei Medianen (entsprechend den Hypothesen H1 bis H9 aus Kapitel 4). - Aufgrund der bisherigen Überlegungen bzw. empirischen Stützen kann eventuell Stufe 3, höchstens aber Stufe 2 erwartet werden.

Als beiläufiger Vorteil des "Verbalisierens" wird die Unterstützung der "symmetriebetonenden" Instruktion erwartet, die ja im Zusammenhang mit den Hypothesen H(akt-4) (S. 148) und H(beob-4) (S. 151) eine bedeutsame Rolle spielt: Die Anweisung, sich später im Prüfdurchgang auch noch an die Rückmeldungen zu erinnern, wird sicherlich nicht dazu führen, sich noch weniger auf f und gleichzeitig um so mehr auf r zu konzentrieren.

Am Anfang dieses Abschnitts war auf die Abhängigkeit theoretischer Aussagen von der jeweils verwendeten experimentellen Methode hingewiesen worden. Es wurde dabei besonders an das Problem gedacht, ob Wiederholungs- oder Behaltens-(Erinnerungs-)Daten (repetition- vs. recall measures) die adäquaten abhängigen Lernmaße zur Überprüfung des Effekt-Gesetzes von THORNDIKE seien (s. dazu NUTTIN, 1953; GREENWALD, 1966; POSTMAN, 1966; s. auch BLÖSCHL, 1969, S. 133).

Eine sehr differenzierte Analyse dieser Kontroverse gibt GREENWALD (s. NUTTIN u. GREENWALD, 1968, S. 148-152). Zunächst stellt er fest, daß hinsichtlich möglicher Erinnerungsdaten (es wird nur das Behalten von R diskutiert) nur die "cued recall"- und nicht die "free recall"-Prozedur in Frage kommen kann (in Übereinstimmung mit THORNDIKE, 1932, S. 458f.). Schließlich kommt er zur Folgerung, daß Erinnerungsdaten, mit gewissen Einschränkungen, den Wiederholungsdaten vorzuziehen seien hinsichtlich der Überprüfung der Annahme von der automatischen Verstärkung durch Belohnungen.

Falls der Vf. die Argumentation GREENWALDs richtig verstanden

hat, hebt das hier vorgeschlagene Effizienzmaß $re(r!)$ das oben besprochene Problem überhaupt auf (jedenfalls im Rahmen der vorliegenden Arbeit, also unter "Verbalisierungs"-Bedingungen).

Da zur Bestimmung von $re(r!)$ ausschließlich die Verbalisierungsdaten "R-R!r!" und " $\bar{R}$ -R!f!" benutzt werden, kann der Erinnerung- nicht vom Wiederholungsaspekt getrennt werden: Aus den Erinnerungsdaten R!r! bzw. R!f! folgt eindeutig die (Nicht-)Wiederholung von R. Es ist hier also "Behalten" und "Leistung" gleichzusetzen (s. BLÖSCHL, 1969, S. 126f.). - Schwierigkeiten tauchen erst auf, wenn auch Vp-Antworten mit unvollständiger Erinnerung miteinbezogen würden (s. insbesondere den g"-Mechanismus bei Vergessen von O, der hinsichtlich des Problems der Wirksamkeit von r und f als problemfremd zu betrachten ist).

7. Hauptuntersuchung

Die in Kapitel 7 zu referierenden Experimente III bis V unterscheiden sich von den Experimenten I und II, abgesehen von den jeweiligen Fragestellungen, im wesentlichen durch folgende äußere Abänderungen:

Die Untersuchungen fanden in einem anderen Versuchsraum statt, in dem die Latenzzeiten mit Hilfe eines Mikrophons und einer elektronischen Stoppuhr gemessen werden konnten. - Die Vpn lernten hauptsächlich "aktiv", an zentraler Stelle allerdings auch "beobachtend". - Es gab jeweils nur einen Lerndurchgang. - Das Lernmaterial, im wesentlichen das gleiche wie in den Voruntersuchungen, wurde mit Hilfe eines Dia-Projektors angeboten; es entfiel also das mühsame manuelle Wechseln der Lernkärtchen. - Außer unter einigen Spezialbedingungen war Rehearsal im allgemeinen gestattet; das Lernen im Metronomrhythmus wurde nicht mehr angewandt. - In den meisten Gruppen wurde im Prüfdurchgang "Verbalisieren" gefordert. - Die Hauptinstruktion war noch stärker "symmetriebe-tonend" formuliert als die zu Experiment II; die Rückmeldung b wurde sorgfältiger eingeführt.

In Experiment III, dem Hauptexperiment, wird versucht werden, alle Fragestellungen (1) bis (5) mit Ausnahme von (2') zu beantworten. Die Experimente IV und V werden sich nur mit Ergänzungsfragen befassen.

7.1 Experiment III: Hauptexperiment

7.1.1 Fragestellung und Hypothesen

Die Fragestellungen (1) bis (3) werden am Ende des Abschnitts behandelt, da erst das Problem der sicheren Ausgangsbasis in Hinblick auf Latenzzeituntersuchungen erörtert werden soll.

Da Fragestellung (5) eine Präzisierung von (4) darstellt,

wird sie jetzt allein besprochen. Aufgrund des Anstoßes durch die Ergebnisse des Experiments II und der daraus folgenden Diskussion eines automatisierten Kontrollprozesses folgte für die Fragestellung (5b) die vorläufige Beantwortung, daß die rein kognitive Informationshypothese nach BUCHWALD (1962) unter gewissen Umständen nicht anwendbar sei. Aus der Diskussion ergab sich die weitere Vermutung, daß man entsprechend Fragestellung (5a) symmetrische bzw. asymmetrische Versuchssituationen nicht nur durch Wechsel der Aufgabenart (z.B. $k=2$ -vs. $k>2$ -Aufgaben) erzeugen könnte, sondern auch durch Wechsel der Lernmodalität ("aktiv" vs. "beobachtend"). Auf die Vorteile dieser letzteren Möglichkeit wurde auf S. 151f. hingewiesen.

Eine weitergehende Beantwortung der Fragestellung (5b) hängt demnach vom Ausgang der Fragestellung (5a) ab.

Für die Fragestellung (5a) lassen sich entsprechend den Ausführungen in 6.3.3 zunächst vier Hypothesen aufstellen. Es werden dabei symmetriebetonende Instruktion, Verbalisieren und Erlaubnis von Rehearsal vorausgesetzt. "Verbalisieren" und "Rehearsal" werden nicht gesondert in die Kurzbezeichnungen der Hypothesen einbezogen.

Hypothese H(akt-4):

Wenn "aktiv" in einer $k=4$ -Situation gelernt wird, ergibt sich Asymmetrie zugunsten von r , d.h. also $re(r!) > 50\%$ (s.S. 148).

Hypothese H(beob-4):

Wenn "beobachtend" in einer $k=4$ -Situation gelernt wird, ergibt sich Symmetrie von r und f , d.h. also $re(r!) = 50\%$ (s. S. 151).

Hypothesen H(akt-2) bzw. H(beob-2):

Wenn "aktiv" bzw. "beobachtend" in einer $k=2$ -Situation gelernt wird, ergibt sich Symmetrie von r und f , d.h. also $re(r!) = 50\%$ (s. S. 137f. u. 150).

Die zu diesen vier Hypothesen gehörenden Versuchsgruppen bilden den Hauptkern des Versuchsplans in Experiment III. Die

Hypothesen (bzw. Versuchsgruppen) sollen für die weitere Diskussion in einem Schema übersichtlich dargestellt werden.

	k = 2	k = 4
"aktives" Lernen	H(akt-2): Symmetrie	H(akt-4): Asymmetrie
"beobachtendes" Lernen	H(beob-2): Symmetrie	H(beob-4): Symmetrie

Es fällt auf, daß nur für eine Versuchsgruppe Asymmetrie vorhergesagt wird, für drei Gruppen dagegen Symmetrie. Nach Fragestellung (5a) könnte man eigentlich mehr Ausgewogenheit erwarten. Die drei hypothetisch als symmetrisch angenommenen Gruppen haben gewisse Kontrollfunktionen, die schon an verschiedenen Stellen angesprochen wurden.

In Anbetracht der relativ starken Abweichungen der experimentellen Durchführung von z.B. den BUCHWALDSchen Experimenten ist es wünschenswert, die Methode des "Verbalisierens", insbesondere auch das Effizienzmaß $re(r!)$, innerhalb eines überschaubaren Rahmens zu erproben.

Die Besonderheit der k=2-Gruppen (nämlich die relativ problemlose Anwendungsmöglichkeit der rein kognitiven Informationshypothese) bringt es mit sich, daß diese Gruppen eine Reihe von Kontrollfunktionen übernehmen können, falls sich H(akt-2) und H(beob-2) bestätigen lassen:

1. Erprobung des neuen Effizienzmaßes; es sollte sich deutlich $re(r!) \approx 50\%$ ergeben. Die erzielte "Deutlichkeit" wäre eine Art Norm für das, was in der H(beob-4)-Gruppe überhaupt an "Symmetrie" erreicht werden könnte.
2. Der Übergang "aktiv $\rightarrow$ beobachtend" geht der Erwartung nach ohne gleichzeitigen Wechsel in der Effizienz von r und f vor sich (s. S. 152f.). Dadurch kann der Übergang "H(akt-4) $\rightarrow$ H(beob-4)" kontrolliert, eventuell relativiert werden.
3. Falls in der H(akt-4)-Gruppe Asymmetrie zugunsten von r auftreten sollte, kann dafür nicht das Effekt-Gesetz nach THORNDIKE verantwortlich sein, wenn gleichzeitig in der

H(akt-2)-Gruppe Symmetrie auftritt. Ebenso kann eine Asymmetrie in der H(akt-4)-Gruppe dann auch nicht mehr dadurch erklärt werden, daß "aktive" Lerner im allgemeinen bessere "Eselsbrücken" finden, die sie nach r beibehalten können.

Innerhalb der k=4-Gruppen gibt es zwei Kontrollmöglichkeiten (s. S. 151f.): Kontrolle der Instruktion durch die H(beob-4)-Gruppe und Eliminierung der k-Abhängigkeit der Latenzzeiten beim Wechsel von Asymmetrie zu Symmetrie.

Diese Verflechtung der Gruppen durch gegenseitige Kontrolle entspricht dem Zweck der Fragestellungen (4) und (5), nämlich eine vergleichsweise präzise und gleichzeitig differenzierte Ausgangsbasis für die Beurteilung der Latenzzeiten zu schaffen. "Vergleichsweise" bezieht sich auf die Möglichkeiten, die ein praktikables BM entsprechend Fragestellung (2') hätte bieten können.

Mit der Besprechung des obigen Hypothesenschemas ist zunächst ein Abschluß bezüglich der Fragestellungen (4) und (5) erreicht. Bevor nun auf die Hypothesen zu den Latenzzeiten eingegangen wird, sollen noch einige ergänzende und untergeordnete Fragen ((a) bis (e)) im Bereich der Wiederholungs- und Erinnerungsdaten behandelt werden.

(a) Auf S. 67 war gesagt worden, daß, falls das strenge BM und damit Fragestellung (2') fallengelassen werden sollten, die Modifikationen des BM und das weitere Vorgehen keinen allzu starken Bruch in der Behandlung der Fragestellungen (2') bzw. (2) bewirken dürften. Da nichtsdestoweniger durch die Erlaubnis von Rehearsal in den Hauptgruppen, H(akt-2) bis H(beob-4), eine beachtliche Abweichung von den Voraussetzungen des BM erzeugt wurde, soll eine Versuchsbedingung mitbehandelt werden, unter der Rehearsal möglichst verhindert (zur weiteren Begründung s. S. 134) und im übrigen "aktiv" gelernt wird.

Die allgemeine Voraussage ist, daß die Ergebnisse dieser Non-rehearsal-Bedingung zwischen denen der entsprechenden Rehearsal-Bedingung, d.h. H(akt-2) bzw. H(akt-4), und den Vorher-

sagen aus dem strengen BM liegen werden. Es wird dabei angenommen, daß unter der Nonrehearsal-Bedingung durch keine experimentelle Manipulation die Vollständigkeit der Verhinderung von Rehearsal erreicht werden kann, wie sie für die strikte Gültigkeit des BM vorausgesetzt ist. Insbesondere wird für $k = 4$ und $re(r!)$ vorausgesagt:

$$re(r!)_{\text{Rehearsal}} > re(r!)_{\text{Nonrehearsal}} > re(r!)_{\text{BM}} = 50\%.$$

Das strenge BM sagt ja wegen $U1:c_r=c_f$, $U2:1-u_r=1-u_f$ und $U3$ (s. S. 34) für alle k $re(r!)=50\%$ voraus. (Das ist nicht zu verwechseln mit den Vorhersagen des BM bezüglich der Wiederholungshäufigkeiten.)

Für die $k=2$ -Nonrehearsalgruppe wird Symmetrie (Informationshypothese!), d.h. $re(r!)=50\%$, also kein Unterschied zur $k=2$ -Rehearsalgruppe oder zum BM vorhergesagt.

Da die Verhinderung von Rehearsal für die Vpn wohl eine Erschwernis beim Lernen bedeutet, ist allgemein für die Nonrehearsalgruppen ein geringerer Lernerfolg, d.h. flachere Profile, zu erwarten.

(Auf die "ohne Rehearsal"- bzw. "Nonrehearsal"-Gruppen wird später unter der Abkürzung "nreh" Bezug genommen.)

(b) Für alle bisherigen Bedingungen, ob $k=2$ oder 4, ob aktiv oder beobachtend, ob Rehearsal oder Nonrehearsal, war "Verbalisieren" im Prüfdurchgang vorausgesetzt worden. Auf die dadurch eventuell eingeschränkte Verallgemeinerbarkeit der Befunde wurde auf S. 159f. (Punkt 2) hingewiesen. Daraus ergibt sich eine weitere ergänzende Fragestellung: Welchen eventuell modifizierenden Einfluß hat "Verbalisieren" auf rr-bb-ff-Profile und Latenzzeiten, verglichen mit unter "Nichtverbalisieren" erhobenen Daten?

(Unter "Nichtverbalisierungs"-Bedingungen kann das Effizienzmaß $re(r!)$ nicht berechnet werden, natürlich aber $re(b)$.)

Zu dem eventuell zu erwartenden Unterschied bei den Profilen kann keine begründete Hypothese aufgestellt werden; die Verbalisierungsanweisung kann sowohl verwirren und behindern als auch den Lernvorgang intensivieren.

Der für das Latenzzeitmuster beim Übergang vom Nichtverbalisieren zum Verbalisieren zu erwartende Effekt wurde auf S. 160f. beschrieben: "Zeitlupeneffekt".

(Auf die Nichtverbalisierungsgruppen wird später unter der Abkürzung "nver" Bezug genommen.)

(c) Es soll nun auf die bisher vernachlässigte Unabhängigkeitsannahme U3 eingegangen werden, obwohl sie jetzt nicht mehr die wichtige Rolle wie im strengen BM spielt (s. dazu S. 34 und 37). Aus den Erinnerungsdaten ist für jedes einzelne Item ersichtlich, ob die zugehörigen Antworten R und Rückmeldungen O gemeinsam oder teilweise oder überhaupt nicht erinnert wurden. Als Maß für das Gemeinsamerinnertwerden kann dann ein Vierfelderkoefizient (z.B. ϕ (= Phi), s. CLAUSS u. EBNER, 1969, S. 256) bestimmt werden.

Bei vollständiger Verhinderung von Rehearsal wird nach U3 $\phi = 0$ erwartet (BM).

Es ist plausibel anzunehmen, daß bei mehr oder weniger erlaubttem Rehearsal ein mehr oder weniger starker (positiver) Zusammenhang zwischen den R- und O-Erinnerungen besteht, also:

$$\phi_{\text{Rehearsal}} > \phi_{\text{Nonrehearsal}} > \phi_{\text{BM}} = 0.$$

Hinsichtlich des Unterschieds zwischen Symmetrie- und Asymmetriegruppen wird die weitere Hypothese aufgestellt:

In den Symmetriegruppen (akt-2; beob-2; beob-4) sind die ϕ -Koeffizienten für r- und f-Items gleich, also $\phi_r = \phi_f$.

In der Asymmetriegruppe (akt-4) ist $\phi_r > \phi_f$.

(Die letztere vermutete Beziehung scheint plausibel in Hinblick auf den postulierten automatisierten Kontrollprozeß. Weitere Hypothesen bezüglich ϕ sind nicht gut zu begründen; $\phi_b < \phi_r$ und ϕ_f zu vermuten ist wohl trivial.)

Die Hypothesen bezüglich ϕ_r und ϕ_f dienen zur weiteren Beleuchtung der Fragestellung (5a). Wie man sich anhand der Darstellung zur statistischen Auswertung (s. S. 185) leicht überlegen kann, geben $re(r!)$ und ϕ_r bzw. ϕ_f voneinander unabhängige Informationen bezüglich Symmetrie oder Asymmetrie.

Die zusätzliche getrennte Auswertung der Erinnerungsdaten als R- bzw. O-recall-Daten dagegen ist redundant, wenn auch unter Umständen illustrativ.

In der späteren Auswertung und Diskussion wird allerdings $re(r!)$ gegenüber $\emptyset$ bevorzugt, weil $re(r!)$ sich besser in den Rahmen der bisherigen Diskussionen einpaßt (s. S. 161f.) und problemloser zu interpretieren ist als $\emptyset$. Außerdem ist $re(r!)$ leicht zu berechnen und, als statistisches Maß gesehen, einwandfrei zu testen. Bei $\emptyset$ muß wegen der Differenzen- und Quotientenbildung mit stärkeren Zufallsschwankungen gerechnet werden.

(d) Aufgrund der Kritik an der Rückmeldung b (s. 6.2.3) werden die Ergebnisse aus den b-Items nur am Rande aufgeführt und interpretiert werden. Auf jeden Fall ist b in den folgenden Experimenten noch mitgeführt worden, um die allgemeine Vergleichbarkeit mit den bisherigen Untersuchungen zu gewährleisten und um speziell $re(b)$ und $re(r!)$ miteinander vergleichen zu können. Aus den bisherigen Überlegungen ist z.B. für die $k=2$ -Gruppen $re(b) < 50\%$ und $re(r!) = 50\%$ zu erwarten.

Der Vergleich von $re(b)$ mit $re(r!)$ ist noch in anderer Hinsicht interessant. Auf S. 139 war gesagt worden, daß $re(b)$ von den Zufallsschwankungen dreier Größen (rr, bb, ff) abhängt. Dagegen tragen zur Variation von $re(r!)$ nur die Schwankungen von $n(R-R!r!)$ und $n(R-R!f!)$ bei. Es ist also bei $re(r!)$ eine geringere Variationsbreite als bei $re(b)$ zu erwarten.

Eine Gleichberechtigung von b mit r und f, wie sie nach dem BM zu fordern ist, kann nicht mehr vernünftigerweise angenommen werden, da infolge der stark symmetriebetonenden Instruktion der Akzent auf r und f liegt und b sehr deutlich als "keine Rückmeldung" eingeführt wird.

(e) Auch wenn das strenge BM nicht mehr vorausgesetzt wird, soll doch die Möglichkeit wahrgenommen werden, g' und g'' direkt aus den Verbalisierungsdaten bestimmen zu können (s. S. 155ff.). Da das Konzept, das hinter $g(g' \text{ bzw. } g'')$ steht, unabhängig von den sonstigen Annahmen ($U1$ bis $U3$) des BM ist, hat die Untersuchung dieser Parameter einen Wert für sich.

Insbesondere interessieren die mögliche k-Abhängigkeit von g' (s. S. 139f.) und die (Un-)Gleichheit von g' und g'' (s. S. 30f.).

Es sollen nun Hypothesen zu den Latenzzeiten behandelt werden.

Die Fragestellungen (1) und (2) auf S. 3 bleiben in der alten Form bestehen, ebenso auch die zugehörigen spezielleren Hypothesen H1 bis H9 (s. Kapitel 4). Dieses trifft für alle bisher beschriebenen Versuchsbedingungen zu. Insbesondere sollen H1 bis H9 sowohl für die Verbalisierungs- als auch die Nichtverbalisierungs-Gruppen gelten: Der "Zeitlupeneffekt" müßte sich zumindest darin bemerkbar machen, daß die Hypothesen H1 bis H9 unter beiden Bedingungen einigermaßen gleich gut unterstützt werden.

Die Fragestellung (3) kann natürlich erst behandelt werden, wenn sowohl die Fragestellungen (1) und (2) als auch die Fragestellungen (4) und (5a) positiv beantwortet worden sind: Eine eventuelle Abhängigkeit der Latenzzeiten von der relativen Effizienz kann eben erst analysiert werden, wenn man die (A-)Symmetrie der Experimentalsituationen, in denen die Zeiten gemessen wurden, abgesichert hat. Für diesen *a n g e - n o m m e n* günstigen Fall soll nun eine Hypothese abgeleitet werden.

Es wird von dem Hypothesenschema auf S. 165 ausgegangen, das für den Augenblick also als positiv beantwortet anzusehen ist. Das Standardvorgehen bei der Ableitung der Hypothesen H1 bis H9 in Kapitel 4 bestand darin, spezifische Annahmen über die zu den (c,1-u)-Wegen gehörigen Antwortzeiten zu machen und die beobachtete Gesamtverteilung als eine Mischung aus einem gewissen Anteil von (c,1-u)-Zeiten und einem Rest relativ langsamer GK-Zeiten zu betrachten.

Dieses Vorgehen genügte für die Ableitung von Hypothesen auf ordinalem Niveau (H7 und H8 gehen allerdings leicht über dieses Niveau hinaus). Für die Aufstellung von mehr und genaueren Hypothesen brauchte man präzisere Kenntnisse über die

prozentualen Anteile der $(c, 1-u)$ - bzw. GK-Zeiten (das wäre über das BM möglich gewesen) und auch über die Größenordnungen z.B. der Umschaltzeiten (s. Begründung zu H3 auf S. 71 f.).

Unter diesen Umständen können nur mit Hilfe von zusätzlichen Annahmen Hypothesen zu dem Schema auf S. 165 abgeleitet werden. Dieses soll nun für den theoretisch wichtigsten Übergang $H(\text{akt-4}) \rightarrow H(\text{beob-4})$ geschehen, also beim Übergang von Asymmetrie zu Symmetrie, ungestört von einer eventuellen k-Abhängigkeit der Latenzzeiten (s. dazu S. 151f., Punkt 2).

Dazu soll im weiteren angenommen werden, daß die "sicheren" $(c_f, 1-u_f)$ -Zeiten, trotz des Umschaltverlustes, immer noch relativ kurz sind im Vergleich zu den GK-Zeiten, d.h. als "schnelle" Anteile zu betrachten sind.

Nun gilt für die $H(\text{akt-4})$ -Bedingung $re(r!) > 50\%$ oder anders ausgedrückt (s. S. 158) $n(R-R!r!) > n(\bar{R}-R!f!)$.

Für die $H(\text{beob-4})$ -Bedingung gilt $re(r!) = 50\%$ oder $n(R-R!r!) = n(\bar{R}-R!f!)$.

Die absoluten Häufigkeiten $n(R-R!r!)$ bzw. $n(\bar{R}-R!f!)$ sind nun aber für die jeweiligen Anteile der "sicheren" $(c_r, 1-u_r)$ - bzw. $(c_f, 1-u_f)$ -Zeiten an den T_{rr} - bzw. T_{ff} -Verteilungen verantwortlich. Diese Anteile, zusammen mit den GK-Zeiten, bestimmen die Lage von T_{rr} bzw. T_{ff} .

Daraus folgt, daß unter der asymmetrischen akt-4-Bedingung zum Median T_{rr} relativ mehr "sichere" $(c_r, 1-u_r)$ -Zeiten und zum Median T_{ff} relativ weniger "sichere" $(c_f, 1-u_f)$ -Zeiten beitragen, verglichen mit der symmetrischen beob-4-Bedingung. D.h., unter der asymmetrischen Bedingung wird die Differenz zwischen T_{ff} und T_{rr} größer sein als unter der symmetrischen Bedingung.

Oder kürzer ausgedrückt (H10 schließt an H9 in Kapitel 4 an):

Hypothese H10: $(T_{ff} - T_{rr})_{\text{sym}} < (T_{ff} - T_{rr})_{\text{asym}}$

Daß eine Hypothese nur für die Differenz $T_{ff} - T_{rr}$ und nicht für die Mediane T_{rr} bzw. T_{ff} einzeln abgeleitet wurde, liegt daran, daß im letzteren Fall besser begründete a-priori-

Annahmen über die Verhältnisse von $n(R-R!r!)$ unter symmetrischen und asymmetrischen Bedingungen hätten vorliegen müssen; das gilt ebenso für $n(\bar{R}-R!f!)$.

Es ist noch einmal daran zu erinnern, daß H10 sich auf den "reinen", d.h. von k ungestörten Übergang von Asymmetrie zu Symmetrie, $H(\text{akt-4}) \rightarrow H(\text{beob-4})$, bezieht. Falls also, wenn überhaupt, ein sehr deutlicher Effekt entsprechend H10 auftreten sollte, besteht die Hoffnung, die Differenz $T_{ff} - T_{rr}$ allgemein als Maß für die relative Effizienz von r und f ausarbeiten zu können.

Weitere Hypothesen zur Fragestellung (3) sollen nicht aufgestellt werden, da sie, wie oben erläutert, nur sehr spekulativ begründet werden könnten. Es entspricht außerdem durchaus der Fragestellung (3), sie a posteriori zu behandeln, d.h. also durch Inspektion der Latenzzeitmuster, die für die vier Bedingungen des Hypothesenschemas auf S. 165 erhalten wurden. Selbstverständlich ist diese Inspektion durch die bisherigen Überlegungen zu den Fragestellungen (1), (2), (4) und (5) zu leiten.

Durch die Anlage des Hypothesenschemas (Hauptversuchsplan) sind prinzipielle teils schon erwähnte Kontrollmöglichkeiten gegeben, die für die Interpretation nützlich sein können: Der eben ausführlicher besprochene Übergang $H(\text{akt-4}) \rightarrow H(\text{beob-4})$ kann durch den Übergang $H(\text{akt-2}) \rightarrow H(\text{beob-2})$ kontrolliert werden, der den Einfluß des Lernmodus "aktiv vs. beobachtend" relativ rein zeigt, da dabei, der Erwartung nach, kein Wechsel der Symmetrieverhältnisse stattfindet.

Aus $H(\text{beob-2}) \rightarrow H(\text{beob-4})$ kann der Einfluß von k ungestört von Symmetriewechsel und Lernmodus abgeschätzt werden. Diesen Einfluß zu kennen ist wichtig für den Vergleich der oben aufeinander bezogenen Übergänge $H(\text{akt-2}) \rightarrow H(\text{beob-2})$ und $H(\text{akt-4}) \rightarrow H(\text{beob-4})$.

$H(\text{akt-2}) \rightarrow H(\text{akt-4})$ beinhaltet sowohl einen Wechsel von k als auch einen von Symmetrie zu Asymmetrie und stellt daher das schwächste Glied innerhalb der Kontrollmöglichkeiten dar.

7.1.2 Methode

Versuchsplan

Der Versuchsplan bestand aus dem schon auf S. 165 dargestellten Hauptkern (Hypothesenschema) und den ergänzenden Versuchsgruppen, in denen entweder versucht wurde, Rehearsal zu minimieren ("nreh"-Gruppen von Nonrehearsal) oder ohne Verbalisieren geprüft wurde ("nver"-Gruppen von Nichtverbalisieren):

	k=2	k=4	k=2	k=4	k=2	k=4
<u>aktiv:</u> →	nreh-2 N=11 5m, 6w	nreh-4 N=12 7m, 5w	akt-2 N=14 7m, 7w	akt-4 N=20 10m, 10w	nver-2 N=13 7m, 6w	nver-4 N=13 6m, 7w
	mit Verbalisieren		beob-2 N=15 7m, 8w	beob-4 N=15 8m, 7w	mit Rehearsal	
<u>beobachtend:</u> →			mit Verbalisieren mit Rehearsal			

Die nreh- bzw. nver-Gruppen sind also nur mit den akt-Gruppen zu vergleichen; untereinander oder mit den beob-Gruppen sind sie nicht vergleichbar.

In den Zellen des Versuchsplans sind die Besetzungszahlen N der Versuchsgruppen eingetragen, aufgeteilt nach männlichen (m) und weiblichen (w) Vpn. Aus technischen Gründen, z.B. Schwierigkeiten bei der Anwerbung von Vpn, wurde die Variable "Geschlecht der Vpn" nur angenähert ausbalanciert. Die theoretisch wichtigsten Gruppen, z.B. akt-4, wurden am stärksten, die weniger wichtigen, z.B. nreh, am geringsten besetzt.

Insgesamt nahmen 131 Vpn teil, von denen 113 im obigen Plan erscheinen. Die Protokolle der übrigen 18 Vpn mußten ausselergiert werden: Bei einigen Vpn hatte die Zeitmessung nicht genügend geklappt, z.B. durch zu leise Stimme oder zu häufiges Räuspern, wodurch das Mikrophon entweder überhaupt nicht oder vorzeitig ansprach. Die anderen Vpn hatten sich nicht an die

Instruktion gehalten, nach der sie im Lerndurchgang ihre Antworten, z.B. 3,5,7 oder 9, möglichst gleich häufig und zufällig abgeben sollten; auf diese Protokolle wird unter 8.4 noch einzugehen sein.

Die Vpn waren überwiegend Studenten (m u. w) aus fast allen Fachbereichen der Universität; die nichtstudentischen Vpn hatten mindestens Mittlere Reife. Bis auf einige Studierende der Psychologie, die "Vp-Stunden" erhielten, nahmen die Vpn freiwillig und unentgeltlich an dem Versuch teil. Das Alter variierte zwischen 19 und 33 Jahren. Keine der Vpn hatte schon bisher Erfahrungen mit Experimenten zur relativen Effizienz von r und f gemacht. - Die Versuche wurden vom Vf. als V1 durchgeführt.

Lernmaterial: Im Verlauf einer Versuchssitzung wurden drei Listen gelernt, eine kurze Einübungsliste a' (5 r-, 5 f- und 6 b-Items) und zwei längere Hauptlernlisten A' und B' (je 13 r-, 13f- und 14 b-Items). Die beiden ersten und letzten Items dienten wiederum als "buffer items". Die mittleren 36 Items (12r:12f:12b) in den Hauptlisten waren, wie in Experiment I, möglichst homogen gemischt. Unter "Item" ist dabei für die "aktiven" Gruppen (akt, nreh und nver) die Kombination "Reizwort plus festgelegte Rückmeldung" zu verstehen, für die beob-Gruppen die Kombination "Reizwort plus festgelegte Rückmeldung und Zahl". Bei den letzteren war auch die Zahlzuordnung mit den Rückmeldungen über die 36 mittleren Positionen ausbalanciert.

Die Listen a', A', B' (s. Anhang, S. 262) waren aus den Reizworten der Listen a, b, A und B des Experiments I erneut sorgfältig zusammengestellt worden, besonders unter dem Gesichtspunkt, keine offensichtlichen "Eselsbrücken" zu provozieren.

Die den Lernlisten a', A' und B' entsprechenden Prüflisten stellten "Teilpermutationen" der Reizwörter aus den Lernlisten dar, d.h.: Die "buffer items" blieben an erster bzw. letzter Stelle; die erste Hälfte der mittleren 36 Items in den Lernlisten tauchte in zufälliger Permutation auch in der

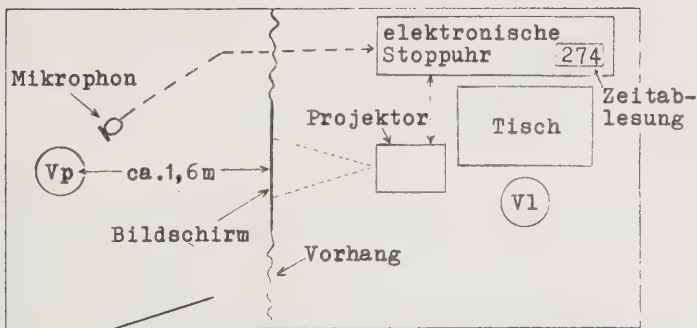
ersten Hälfte der mittleren 36 Reizwörter in den Prüflisten auf; entsprechend geschah es mit der zweiten Hälfte. Das hatte den Zweck, daß zwar Reihenfolgeeffekte möglichst ausgeschaltet waren, daß aber zwischen Lernen eines Items und Prüfung desselben mindestens 18 andere Items vorkamen. Es konnte also nicht passieren, daß ein Item an 37. Position in einer Lernliste schon an 3. Stelle im Prüfdurchgang erschien.

Die Lernlisten wurden innerhalb der jeweiligen Versuchsgruppen der einen Hälfte der Vpn in der Reihenfolge A'B', der anderen Hälfte in der Reihenfolge B'A' angeboten, d.h. die Listen A' bzw. B' kamen je zur Hälfte einmal in Position I bzw. II vor. Die Reihenfolge war mit der Variable "Geschlecht der Vpn" möglichst ausbalanciert.

Als Antwortalternativen wurden wieder die Ziffern 3 und 5 bzw. 3, 5, 7 und 9 benutzt.

Versuchsdurchführung

Die Versuchsanordnung sei durch folgende Skizze veranschaulicht:



Der Versuchsraum hatte keine Fenster, nur einen Ventilator, der aber bei der Versuchsdurchführung nicht eingeschaltet war. Während des Versuchs wurde die Deckenbeleuchtung ausgeschaltet, so daß nur der Tisch, auf dem der V1 protokollierte, von einer Tischlampe beleuchtet war; dabei wurde der schwarze Vorhang ganz zugezogen.

Die elektronische Stoppuhr war in folgender Weise mit dem Projektor gekoppelt¹⁾: Wenn der V1 auf den Betätigungsknopf für den Projektor drückte, wurde ein Dia, auf dem z.B. ein Reizwort stand, aus dem Magazin geholt und auf den Bildschirm (durchleuchtende Mattscheibe) projiziert. Unmittelbar mit der Projektion gekoppelt wurde die Stoppuhr automatisch gestartet. Sobald die Vp antwortete, wurde über das Mikrophon die Uhr gestoppt, und die gemessene Zeit erschien (auf 1/100 sec genau) in Leuchtziffern im Anzeigefenster der Stoppuhr. Die Zeitanzeige blieb bis zum Start des nächsten Dia stehen, so daß der V1 bequem die Antwortzeit protokollieren konnte.

Da der Projektor vom V1 nur weiterzuschalten war, wenn das Mikrophon angesprochen hatte, war es wichtig, daß die Vpn laut genug antworteten.

Das Dia verschwand gleichzeitig mit dem Ansprechen des Mikrophons. Es blieb auf keinen Fall länger als 1,8 sec stehen; den Vpn wurde gesagt, daß dieses keinen Zeitdruck bedeuten sollte.

Das Schriftbild auf dem Bildschirm hatte das Format: 4,5 cm (Höhe) mal maximal 19 cm (Breite). Die Dias waren mit schwarzer Tusche beschriftet: die Reizworte in großen Druckbuchstaben, die Rückmeldungen "richtig" und "falsch" für die beobachteten Gruppen in kleinen Druckbuchstaben, die Rückmeldung b als Bindestrich "—".

Anders als in Experiment I wurde unter Beobachtungsbedingungen ein Item nicht auf einem Dia (bzw. Kärtchen), sondern auf drei Dias angeboten, z.B.: <FLUT> < 7 > <falsch>. Dieses hatte den Zweck, die Ähnlichkeit, zumindest in zeitlicher Hinsicht, mit den "aktiven" Bedingungen zu erhöhen. - Es war in 5.2.4 kritisiert worden, daß bei Darbietung eines Items auf einem Kärtchen die Vp sofort sehen könnte, um was für eine Art Item (r, f oder b) es sich gerade handele. In Experiment I war diese Möglichkeit allerdings noch schwerwiegender,

1) Ich danke Herrn Professor K.-Th. Kalveram für den Aufbau der elektronischen Zeitmeßapparatur.

da sie die strenge Unterbindung von Rehearsal unterlief.

Die Dias für die Lern- und Prüfdurchgänge standen in Magazinen neben dem Projektor. Für einen Lerndurchgang unter Beobachtungsbedingungen wurden $3 \times 40 = 120$ Dias benötigt (40 Items bestehend aus Reizwort, Zahlalternative, Rückmeldung). Diese 120 Dias waren in drei Magazinen untergebracht, die während des Lerndurchgangs vom V1 möglichst ohne merkbare Unterbrechung nacheinander in den Projektortransportmechanismus eingeschoben wurden.

Über dem Bildschirm war ein Kärtchen befestigt, das mit den jeweiligen Antwortalternativen beschriftet war, also mit 3,5 oder mit 3,5,7,9. Dieses hatte den Zweck, die Vpn jederzeit an ihre möglichen Antworten zu erinnern, eventuell im Lerndurchgang die Gleichhäufigkeit der Zahlen zu unterstützen. Auch bei der mündlichen Erläuterung der Instruktion durch den V1 war das Kärtchen nützlich.

Der Versuchsablauf wird zunächst für die akt-4-Gruppe dargestellt (die Abänderungen für die übrigen Gruppen werden anschließend beschrieben):

Zu Beginn erhielt die Vp die Instruktion (s. Anhang, S.266f.) mit der Aufforderung, sie sorgfältig durchzulesen. Das Durchlesen dauerte ca. 5 min. Der V1 saß so lange im Hintergrund an seinem Tisch. Anschließend beantwortete der V1 etwaige Fragen zum Verständnis der Instruktion und ließ sich dann den Versuchsablauf noch einmal in großen Zügen von der Vp beschreiben. Gegebenenfalls wies der V1 ein wiederholtes Mal auf die gleichmäßige Beachtung von r und f hin.

Dann kündigte der V1 eine kurze Einübungsliste an, schaltete die Deckenbeleuchtung aus, zog den Vorhang hinter sich zu und begann mit der Projektion der Einübungsliste a'. Falls die Vp zu leise Antworten gab, so daß das Mikrophon nicht ansprach, wurde sie gebeten, lauter zu sprechen.

Unmittelbar nach jeder Antwort der Vp sagte der V1 laut und mit möglichst gleichbleibendem Tonfall die für das jeweilige Item festgelegte Rückmeldung. Dann notierte er die Antwort

der Vp, las die gemessene Zeit ab und notierte sie gleichfalls. Anschließend betätigte er wieder den Projektorknopf zur Darbietung des nächsten Dias. Die Zeit zwischen Rückmeldung und Betätigung des Projektorknopfes betrug ca. 5 sec. (Diese Zeit konnte von der Vp für Rehearsal benutzt werden.)

Die Antwortzeiten im Lerndurchgang wurden zwar nur am Rande ausgewertet (hier nicht dargestellt), das Ablesen und Notieren aber hatten die wichtige Funktion, den zeitlichen Ablauf einigermaßen unter Kontrolle zu bringen. Dieser zeitliche Ablauf war in Hinblick auf die veränderten Bedingungen in den nreh- und beob-Gruppen zu beachten.

Nach dem letzten Item der Liste a' kündigte der V1 den Prüfdurchgang an und wechselte dabei das Magazin. Die Zeit zwischen Lern- und Prüfdurchgang betrug ca. 30 sec.

Im Prüfdurchgang projizierte der V1 die Reizwörter und notierte die Antworten der Vp und die Antwortzeiten. Hier vermied es der V1, einen besonders straffen Zeitablauf einzuhalten, um die Vp nicht in einen bestimmten Rhythmus zu drängen.

Nach dem 8. Item im Prüfdurchgang unterbrach der V1 den Ablauf und gab eine mündliche Zusatzinstruktion (Einführung von "Verbalisieren"):

"Wir haben jetzt gerade die Hälfte des Prüfdurchgangs beendet. - Es kommt nun noch eine kleine Abänderung in der Instruktion. - Wenn Du auf das Reizwort hin Deine Zahl gesagt hast, wie Du es eben gerade gemacht hattest, dann sage, bitte, noch anschließend, wie Du zu dieser Zahl gekommen bist. - D.h., versuche Dich zu erinnern, welche Zahl Du im Lerndurchgang genannt hast und welche Rückmeldung ich darauf gegeben habe. -

Drücke dabei die Sicherheit Deiner Erinnerungen an den Lerndurchgang in drei Stufen aus, nämlich:

sicher - unsicher - vergessen

Also: sicher bzw. unsicher, wenn Du Dich überhaupt erinnerst; - und vergessen bei völligem Fehlen irgendeiner Erinnerung. - Die Reihenfolge ist im weiteren Prüfdurchgang also folgende:

1. Reizwort

2. Deine endgültige Zahl, wie Du es eben schon gemacht hattest.

3. Deine Erinnerung an die Zahl, die Du im Lerndurchgang genannt hattest, mit der entsprechenden Sicherheitsstufe.

4. Deine Erinnerung an meine Rückmeldung, mit der entsprechenden Sicherheitsstufe.

Wir fahren jetzt im Prüfdurchgang fort. Ich werde Dir bei der Einübung dieser ziemlich schwierigen Antwortweise einhelfen."

Im weiteren Verlauf des Prüfdurchgangs notierte der V1 dann noch zusätzlich die Erinnerungsantworten mit den Konfidenz-urteilen. Die Schwierigkeiten der Vpn beim "Verbalisieren" in dieser Einübungsphase waren im allgemeinen erheblich und die Hilfen des V1 unerlässlich.

Nach Beendigung des Prüfdurchgangs zur Liste a' kam der V1 wieder nach vorne zu der Vp, wies sie darauf hin, daß dieses ja noch die Einübungsphase wäre und daß alle Vpn hier Schwierigkeiten hätten. Es wurden etwa neu aufgetretene Fragen zur Hauptinstruktion beantwortet. Dann betonte der V1, daß die Vp von jetzt ab im Prüfdurchgang immer die Erinnerungen an den Lerndurchgang mit den Sicherheitsurteilen zusätzlich angeben solle, und erklärte die Verbalisierungsanweisung noch einmal anhand einer Karte mit Beispielen. Die Karte sah so aus:

Wie sicher sind Deine Erinnerungen an Deine Zahl und meine Rückmeldung im Lerndurchgang?
sicher oder unsicher oder vergessen

Beispiele für das Antworten im Prüfdurchgang:

WILD	5	5: unsicher	richtig: sicher
RAUB	3	5: sicher	falsch: unsicher
STIL	3	vergessen	blank: sicher
FLUT	5	vergessen	vergessen

Falls die Vp keine Fragen mehr hatte, wies der V1 noch einmal auf die Lernaufgabe in der Hauptinstruktion hin: Die Vp solle sich bemühen, nach r die Zahlen zu wiederholen, nach f die Zahlen nicht mehr zu wiederholen; beides stelle einen Lernerfolg dar; nach b wäre kein Lernerfolg möglich. (Damit war die Einübungsphase beendet.)

Dann kündigte der V1 die nächste Lernliste an und ging wieder zu seinem Tisch hinter den Vorhang zurück. Es wurden nun der Lern- und der Prüfdurchgang mit der Hauptliste A' (bzw. B'), wie bei Liste a' beschrieben, durchgeführt (Übungsstadium: Position I). Im Prüfdurchgang mußte die Vp von Anfang an "verbalisieren".

Anschließend war eine Pause von ca. 5 min, in der sich der

V1 mit der Vp außerhalb des Versuchsraums aufhielt und mit ihr über Themen sprach, die nichts mit dem Versuch zu tun hatten. Das diente primär zur Erholung, aber auch dazu, daß die Vp sich nicht mehr zu sehr mit der gerade beendeten Liste beschäftigte und eventuell sich globale Strategien überlegte, die den Versuchszweck unterlaufen hätten (s. dazu unter 8.4.).

Nach der Pause wurden Lern- und Prüfdurchgang mit der zweiten Hauptlernliste B' (bzw. A') durchgeführt (Übungsstadium: Position II).

Nach Abschluß dieser Lernliste stellte der V1 im postexperimentellen Interview folgende Fragen:

F1: Haben die Rückmeldungen r bzw. f auf Dich irgendwie emotional gewirkt?

F2: Hat b auf Dich emotional gewirkt?

F3: Hast Du Dich besonders auf r-Wörter konzentriert?

F4: Hast Du versucht, b-Wörter zu lernen?

F5: Beschreibe, bitte, Deine allgemeine Lernstrategie!

(Diese Fragen wurden notfalls vom V1 noch näher erläutert.)

Die Versuchsdauer schwankte zwischen 95 und 105 min. Auf Wunsch der Vpn, und das war meistens der Fall, zählte der V1 den Lernerfolg aus und erklärte den Versuchszweck.

Damit ist der vollständige Versuchsablauf für die akt-4-Gruppe beschrieben.

Für die akt-2-Gruppe war nur die Hauptinstruktion etwas abgeändert: Statt der Zahlen 3,5,7 und 9 wurden als Antwortmöglichkeiten nur noch 3 und 5 angegeben.

Für die nver-Gruppen entfiel die Zusatzinstruktion, mit der in den anderen Gruppen "Verbalisieren" eingeführt wurde. Die Versuchsdauer verkürzte sich auf ca. 85 min.

Bei den nreh-Gruppen wurde versucht, den zeitlichen Ablauf im Lerndurchgang möglichst gleich wie bei den akt-Gruppen zu gestalten, bis eben auf die versuchte Minimierung der Zeit zwischen Rückmeldung des V1 und Erscheinen des nächsten Reizwortes auf dem Bildschirm. Die Abfolge für ein Item war in den

akt- bzw. nreh-Gruppen folgende (das Schriftbild entspricht angenähert der verstreichenden Zeit):

akt:

<WIND>..5;richtig;5 und Zeit notieren; Druck ; <GOLD> usw.

nreh:

<WIND>..5;5 und Zeit notieren;richtig; Druck ; <GOLD> usw.
still richtig

Eckige Klammern bedeuten projizierte Dias. "5" wird von der Vp, "richtig" vom V1 ausgesprochen. "Druck" bedeutet Betätigung des Projektorknopfes und Transport des Dias bis zum Erscheinen von <GOLD> auf dem Bildschirm.

"richtig
still"} weist auf stilles Aussprechen von "richtig" durch den V1 hin.

"Druck
richtig"} bedeutet Betätigung des Projektorknopfes und lautes "richtig" vom V1 während des Dia-Transports. Unmittelbar danach erscheint dann <GOLD>.

Die Zeit für "5 und Zeit notieren" betrug ca. 5 sec.

Die je nach den Rehearsal-Bedingungen unterschiedliche zeitliche Abfolge in der Betätigung des Projektorknopfes und der Rückmeldung war vom V1 in den Vorversuchen eingeübt worden.

Auch für die beob-Gruppen wurde versucht, den zeitlichen Ablauf im Lerndurchgang mit dem in den akt-Gruppen vergleichbar zu machen. Das war allerdings nur mit Einschränkungen möglich: 1. Der Transportmechanismus erforderte eine gewisse Zeit; 2. "aktiven" Vpn stand es frei, wie schnell sie im Lerndurchgang auf das Reizwort reagieren wollten; daraus resultierten starke individuelle Schwankungen in den Antwortzeiten. - Der zeitliche Ablauf war für eine beob-Vp folgender:

<WIND>Druck< 5 >Druck<richtig>..5sec..Druck<GOLD> usw.

Eckige Klammern bedeuten wieder Dias. Die Vpn hatten die Anweisung, die Reizworte, Zahlen und Rückmeldungen laut mitzulesen (< — > war als "blank" auszusprechen). Die ..5sec.. waren für Rehearsal vorgesehen, wie auch in den anderen Gruppen (akt, nver) mit Rehearsalerlaubnis.

Der allgemeine Versuchsablauf unter beob-Bedingungen begann zunächst wie bei den anderen Gruppen mit dem Lesen der Hauptinstruktion, Fragen an den V1 und der kurzen Darstellung der Lernaufgabe durch die Vp.

Dann erklärte der V1 der Vp, daß sie an einer Versuchsbedingung teilnähme, die etwas von der vorliegenden Instruktion abweiche. Anhand der Beispiele in der Hauptinstruktion wurde dann "beobachtendes Lernen" eingeführt: Die Vp möge sich die Beispiele für den Lerndurchgang als das Lernprotokoll einer anderen Versuchsperson vorstellen. Sie könne, ebenso wie diese andere Versuchsperson, beispielsweise aus der Sequenz FLUT 5 falsch entnehmen, daß sie im Prüfdurchgang z.B. nicht mehr mit 5 antworten dürfe. Der P r ü f durchgang würde jedenfalls g e n a u s o ablaufen wie bei dieser anderen Versuchsperson!

Die Vp erhielt darauf als Einübungsliste für "beobachtendes Lernen" die erste Hälfte der Liste a' (s. Anhang, S. 262). Nach dem Prüfdurchgang gab der V1 als weitere zusätzliche Instruktion die "Verbalisierungs"-Anweisung, die er anhand der Beispiele aus der Hauptinstruktion und der Karte (s. S. 179) mit möglichst denselben Worten wie denen in der Zusatzinstruktion auf S. 178 erklärte. Danach bekam die Vp als Einübungsliste für "Verbalisieren" die zweite Hälfte der Liste a'. - Der V1 gab dann das Ende der Einübungsphase an und beantwortete noch Verständnisfragen der Vp.

Der weitere Versuch lief wie bei den akt-Gruppen ab, bis eben auf den Lernmodus "beobachtend".

Im Vergleich zu der Einführung des Beobachtungslernens in Experiment I (s. Instruktion auf S. 77-80) wurde hier betont versucht, die Einstellung der Vpn zu der Lernaufgabe sowohl bei den akt- als auch bei den beob-Gruppen möglichst aneinander anzugleichen. Dieses Ziel sollte durch die einheitliche Anfangsphase (dieselbe Hauptinstruktion und dieselbe mündliche Kurzdarstellung des Versuchs durch die Vp!) und durch die Einführung des beobachtenden Lernens unmittelbar anhand der

für "aktive Lerner" gültigen Hauptinstruktion angenähert werden.

Die Angleichung der Lerneinstellungen war zu erstreben, damit im Sinne von 6.3.3 die beob-Vpn die $k=4$ -Bedingung als eine Mehrfachwahlsituation auffaßten und demzufolge p r i n z i p i e l l auch bei ihnen der (für $k>2$ postulierte) automatisierte Kontrollprozeß angestoßen werden k o n n t e.

Zusätzlich ist zu betonen, daß die Versuchsdurchführung für die akt- und beob-Gruppen so gestaltet war, daß vom rein kognitiven Informationsstandpunkt her die Lernaufgabe für beide Gruppen die gleiche war!

Zur statistischen Auswertung

Bezüglich des allgemeinen Vorgehens in der Auswertung sei auf 5.2.2 (S. 96ff.) hingewiesen.

Profile werden für alle Gruppen wie bisher aus den Wiederholungsdaten berechnet. Etwaige Vergleiche zwischen Profilen sollen deskriptiv behandelt werden.

Die weitere Auswertung der Verbalisierungsdaten ist näher zu erläutern. Zur Veranschaulichung werden die Daten eines Protokolls (Vp-Nr. M 27) genommen, die zu den 12 r-Items der Liste A' gehören (die r-Items sind dort natürlich mit den f- und b-Items gemischt).

Lerndurchgang		Prüfdurchgang			Verschlüsselung		
S	R O	R-corr	R-recall	O-recall	R-corr	R-recall	O-recall
KNIE	3 r	3	3?	b!	x	R?	?
BROT	3 r	3	5!	b!	x	?	?
FELL	5 r	5	3!	b?	x	?	?
KORB	3 r	3	5!	f!	x	?	?
SAAT	3 r	3	3?	r!	x	R?	r!
KOPF	3 r	3	3!	r?	x	R!	r?
MUND	5 r	5	5!	r!	x	R!	r!
QUAL	3 r	5	3!	f!		R!	?
HAAR	5 r	5	5!	r!	x	R!	r!
ZELT	5 r	5	5!	b?	x	R!	?
DACH	5 r	3	?	?		?	?
HAUS	5 r	3	3?	r?		?	r?

Unter "Lerndurchgang" sind nach den Reizworten S die Antworten R der Vp angegeben, auf die dann die Rückmeldung O=r folgt. Unter "Prüfdurchgang" sind in der ersten Spalte die (Nicht-)Wiederholungsantworten der Vp angegeben, dann in den beiden nächsten Spalten ihre Erinnerungsangaben. Die Konfidenzurteile sind dabei abgekürzt, z.B: "3 sicher" durch "3!", "r unsicher" durch "r?", "vergessen" durch "?" usw.

Unter "Verschlüsselung" sind in der ersten Spalte die Wiederholungen nach r, also Lernerfolge, durch "x" markiert. In der zweiten (dritten) Spalte ist die sichere bzw. unsichere Erinnerung an die Antwort R (Rückmeldung O) aus dem Lerndurchgang durch R! bzw. R? (r! bzw. r?), die vergessene bzw. fälschliche Erinnerung durch "?" verschlüsselt.

Das einfachste ist die übliche Auswertung des von den r-Items beigesteuerten Lernerfolgs aus der ersten verschlüsselten Spalte: $rr=9/12=75\%$. Dieser Wert ergibt also einen Punkt der Profile aus den Wiederholungshäufigkeiten rr, bb und ff.

Die verschlüsselten Erinnerungsdaten (zweite und dritte Spalte) werden in einer 3 x 3-Kontingenztafel zusammengefaßt:

		O-recall			
		r!	r?	?	
R-recall	R!	//	/	//	5
	R?	/		/	2
	?		/	///	5
		3	2	7	12

(Zur Kontrolle können die Randsummen berechnet werden, deren Summe 12 ergeben muß.)

In 6.3.4 auf S. 158 war bei der Definition von $re(r!)$ begründet worden, warum die weiteren Diskussionen und Auswertungen sich auf die "sicher"-Kategorie der Konfidenzurteile beschränken sollten. Die 3 x 3-Kontingenztafel wird daher auf eine Vierfeldertafel reduziert durch Streichung der mittleren Zeile (R?) und Spalte (r?); dieses Verfahren läuft also auf einen "Extremkategorien"-Vergleich hinaus:

		O-recall	
		r!	?
R-recall	R!	2	2
	?	0	4
		2	6
			8

(Wieder kann eine partielle Kontrolle über die Randsummen erfolgen, die jetzt aber nicht mehr 12 ergeben müssen.)

Auf die gleiche Art wurden auch die zu den f- und b-Items gehörenden Erinnerungsdaten in entsprechenden Vierfeldertafeln zusammengefaßt. Diese Tafeln waren der Ausgangspunkt für weitere Auswertungen.

Zur zusätzlichen Illustration sei noch ein Vierfelderschema für f-Items angegeben:

		O-recall	
		f!	?
R-recall	R!	3	1
	?	1	2
		4	3
			7

Aus den obigen zwei Tafeln lassen sich verschiedene inhaltlich relevante Größen berechnen:

1. das auf S. 158 definierte relative Effizienzmaß: $re(r!) = n(R-R!r!)/\langle n(R-R!r!) + n(\bar{R}-R!f!) \rangle = 2/\langle 2+3 \rangle = 40\%$;
2. R-recall: Die relativen Häufigkeiten der sicheren Erinnerungen an R seien durch rr' ! und ff' ! abgekürzt; rr' ! = $4/8 = 50\%$ und ff' ! = $4/7 = 57\%$;
3. O-recall: Die relativen Häufigkeiten der sicheren Erinnerungen an O seien durch rr'' ! und ff'' ! abgekürzt; rr'' ! = $2/8 = 25\%$ und ff'' ! = $4/7 = 57\%$;
4. ein Maß für die Abhängigkeit der Erinnerungen an R und O voneinander: Als Abhängigkeitsmaß wurde der bezüglich der Randfelddbesetzungen korrigierte "interspezifische Assoziationskoeffizient" ϕ (= Phi) nach COLE berechnet (s. CLAUSS u. EBNER, 1970, S. 256, Formel 4.50a). Es ergibt sich z.B. aus der Tafel für die f-Items:

$$\phi_f = (3 \cdot 2 - 1 \cdot 1) / \langle 7 \cdot \min(1, 1) + (3 \cdot 2 - 1 \cdot 1) \rangle = 0,46.$$

ϕ_r wurde nicht als Berechnungsbeispiel gewählt, weil die r-Tafel eine Zelle mit Nullbesetzung enthält; LIENERT (1973, S. 538) empfiehlt hier, in allen Zellen um 1 zu erhöhen.

Das diesbezügliche Problem, d.h. Nullbesetzungen, ergab sich aber bei der tatsächlichen Auswertung der Protokolle nicht, da die Koeffizienten ϕ_r , ϕ_f und ϕ_b aus den Vierfeldertafeln berechnet wurden, die durch Zusammenlegen aller individuellen Tafeln innerhalb einer Gruppe entstanden.

Zum Zweck der deskriptiven Beschreibung des Zusammenhangs zwischen R und O ist dieses Verfahren erlaubt. Soll allerdings z.B. gegen $\phi = 0$ getestet werden, ist das Vorgehen wegen der teilweisen Abhängigkeit der Zellenhäufigkeiten problematisch. (Auf das entsprechende Problem bei den Latenzzeitverteilungen war auf S. 99 hingewiesen worden.)

Für vorliegende Zwecke genügt die deskriptive Funktion des so bestimmten Abhängigkeitsmaßes, da es nicht mehr um den "harten" Test der Unabhängigkeitsannahme U_3 im BM geht (s. auch S. 168).

Diese vier Aspekte, also $re(r!)$, $rr'!$ und $ff'!$, $rr''!$ und $ff''!$, ϕ_r und ϕ_f , nach denen die beiden Vierfeldertafeln ausgewertet werden können, geben keine voneinander unabhängigen Informationen. Die für die vorliegende Diskussion relevanteste Kenngröße ist $re(r!)$. Die Koeffizienten ϕ_r und ϕ_f erfassen einen anderen Aspekt der beiden Tafeln als $re(r!)$ und sollen deshalb später auch gesondert diskutiert werden.

Die R- bzw. O-Erinnerungshäufigkeiten $rr'! \dots ff'!$ geben zu $re(r!)$ und ϕ_r , ϕ_f nur noch stark redundante Informationen. Sie werden aber auch deskriptiv dargestellt werden, um die Ergebnisse mit den Resultaten der R- bzw. O-recall-Bedingungen aus den Experimenten I und II (s. S. 106 und 108) vergleichen zu können; in diesen Experimenten konnte $re(r!)$ noch nicht berechnet werden.

Die vier Hauptthesen (S. 164) werden mit Hilfe des WILCOXON-Tests überprüft (SACHS, 1969, S. 312). Dazu müssen sie etwas umgeschrieben werden. Entsprechend der Definition von

$re(r!)$ auf S. 158 können die Hypothesen auch mit Hilfe von $n(R-R!r!)$ und $n(\bar{R}-R!f!)$ ausgedrückt werden:

$H(akt-4): n(R-R!r!) > n(\bar{R}-R!f!);$

$H(akt-2), H(beob-2) \text{ und } H(beob-4): n(R-R!r!) = n(\bar{R}-R!f!).$

Es werden $H(akt-4)$ einseitig, die anderen Hypothesen zweiseitig getestet. - Die für den Test benötigten Wertepaare ergeben sich für jede Vp aus den beiden Zellen $(R!r!)$ und $(R!f!)$ ihrer Über beide Positionen I und II (d.h. Listen A' und B') aufsummierten r- und f-Vierfeldertafeln.

Ebenso werden auch die beiden nreh-Gruppen gegen $re(r!) = 50\%$ getestet ($k=4$: einseitig; $k=2$: zweiseitig; s. S. 167).

Der Übungseffekt (Position I vs. II) ist hier von keinem besonderen Interesse. Nichtsdestoweniger werden zu Beginn der Auswertung die Profile, die Effizienzmaße $re(b)$ und $re(r!)$ und die Latenzzeitmuster zunächst für die Positionen I und II getrennt berechnet zum Zwecke einer a-posteriori-Inspektion.

Aus den Differenzen (= Variationsbreiten) zwischen den Positionen I und II sowohl von $re(b)$ als auch $re(r!)$ werden die mittleren Variationsbreiten über alle Gruppen gebildet. Nach S. 169 ist für $re(r!)$ eine kleinere Variationsbreite zu erwarten.

Die Schätzungen der Parameter g' bzw. g'' (s. S. 169f.) werden getrennt für $k=2$ und $k=4$ berechnet. Um genügend Daten zu bekommen, wird die Auswertung ausnahmsweise die Konfidenzurteile "sicher" und "unsicher" gemeinsam umfassen. Für die Schätzung von g' werden die Fälle $(R-R,b)$ und $(\bar{R}-R,b)$ gezählt; dann ist $g' = n(R-R,b) / \langle n(R-R,b) + n(\bar{R}-R,b) \rangle$.

Entsprechend ergibt sich: $g'' = n(R-R,?) / \langle n(R-R,?) + n(\bar{R}-R,?) \rangle$. "?" bedeutet, daß die Vp bezüglich der zu erinnernden Rückmeldung explizit "vergessen" sagte.

Zur Kontrolle, wie adäquat die Vpn die beiden Konfidenzurteile "sicher" und "unsicher" benutzt haben, werden die a-posteriori-Wahrscheinlichkeiten für die Erinnerungen an die Rückmeldungen berechnet werden. Eine a-posteriori-Wahrscheinlich-

keit in diesem Zusammenhang ist die Wahrscheinlichkeit, daß eine Erinnerung an O (r,f oder b) korrekt ist, gegeben ein bestimmtes Konfidenzurteil (z.B. MURDOCK, 1970, S. 293). Es wird sich auf die Auswertung der O-Erinnerungen beschränkt, da die Auswertung der "gemischten" Fälle, z.B. (R|r?), (R?f!) usw., schwierig geworden wäre und eine eventuell unterschiedliche Behandlung der Konfidenzstufen in Abhängigkeit von O sich vermutlich deutlicher bei den Erinnerungen an die Rückmeldungen selbst zeigen könnte. Allerdings gibt es a priori keinen Grund, eine derartige unterschiedliche Behandlung zu erwarten.

7.1.3 Ergebnisse

Wiederholungsdaten

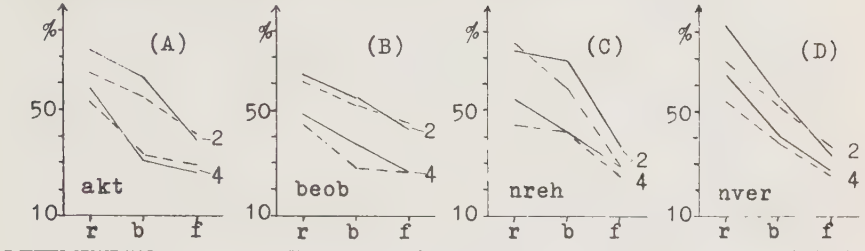
Die Abb. 7.1.3-1 ist so angelegt, daß die jeweiligen Profile für die einzelnen Gruppen (akt, beob, nreh und nver) der besseren Vergleichbarkeit wegen direkt untereinander stehen.

In der oberen Hälfte der Abbildung sind die relativen Wiederholungshäufigkeiten rr, bb und ff in der bekannten Form dargestellt.

Aus der ersten Reihe (A) bis (D) ist der allgemeine Verlauf der Profile in Abhängigkeit vom Übungsstadium (Position I vs. II) zu entnehmen. Die Profile sind regulär, d.h. sie fallen von rr über bb nach ff monoton ab. Der Übungseffekt, wie man ihn erwartet, zeigt sich am deutlichsten bei den akt-Gruppen (A), d.h. in Position II sind rr größer und ff kleiner als in Position I. - Auffällig ist die Umkehr dieser Erwartung für die nreh-2-Profile (C): In Position II ist der Lernerfolg geringer als in Position I. Diese Umkehr ist aber wohl kaum mehr als eine starke Zufallsschwankung zu werten.

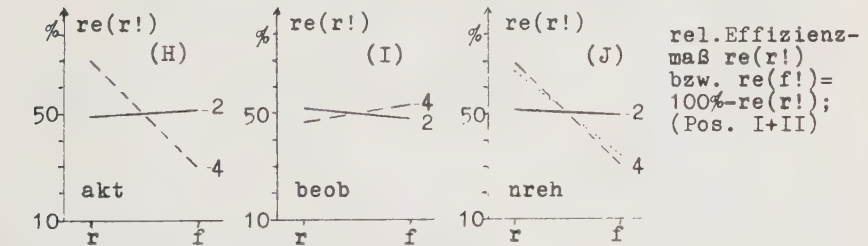
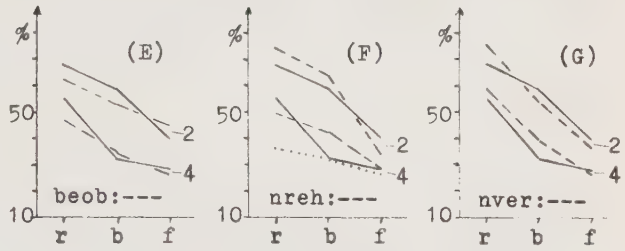
In der zweiten Reihe (E), (F) und (G) sind die, über die Positionen I und II gemittelten, Profile der beob-, nreh- und nver-Gruppen je im Vergleich mit den Profilen der entsprechenden akt-Gruppe dargestellt (s. die zentrale Stellung der

Übungsstadium: Position I (---) vs. Position II (—);
rel.Wied.häuf. rr, bb, ff

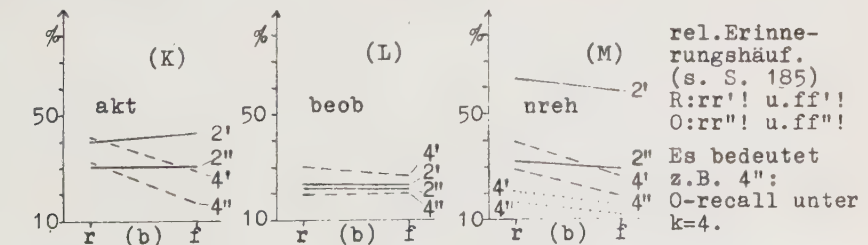


rel.Wied.häuf.
rr, bb, ff

Vergleich der
akt-Gruppen(—)
mit den anderen
Gruppen (---);
(Pos. I+II)



rel.Effizienz-
maß $re(r!)$
bzw. $re(f!)$ =
 $100\% - re(r!)$;
(Pos. I+II)



rel.Erinner-
ungshäuf.
(s. S. 185)
 $R:rr'!$ u. $ff'!$
 $O:rr''!$ u. $ff''!$
Es bedeutet
z.B. $4''$:
O-recall unter
 $k=4$.

Abb. 7.1.3-1: Gesamtübersicht der Wiederholungs- und Erinnerungsdaten (Experiment III).
("nreh-4" aus Experiment IV:)

akt-Bedingung im Versuchsplan auf S. 173).

Die beob-Gruppen haben offensichtlich einen schlechteren Lernerfolg erzielt als die akt-Gruppen. Beiläufig erwähnt sei die Übereinstimmung des trendmäßigen Unterschiedes zwischen der akt-4- und beob-4-Gruppe mit dem von MARX u. WITTER (1972) erhaltenen Ergebnis (s. dort die beiden Abbildungen auf S. 55 und 56; nur der erste Durchgang darf zum Vergleich herangezogen werden).

Die nreh-2-Gruppe zeigt unerwartet einen augenscheinlich besseren Lernerfolg als die akt-2-Gruppe, der Rehearsal erlaubt war. Das nreh-4-Profil ist leicht irregulär, insofern es nicht die typische Form von $k > 2$ -Profilen zeigt. Für beide Abweichungen ist wohl die Zufallserklärung am naheliegendsten.

Wichtiger ist der Vergleich der nver- mit den akt-Profilen. Er gibt eine Teilantwort auf die ergänzende Fragestellung (b) bezüglich des Einflusses von "Verbalisieren" auf Profile und Latenzzeiten (s. S. 167f.). Im ganzen gesehen scheint unter nver-Bedingungen besser gelernt zu werden als unter akt-Bedingungen mit gefordertem "Verbalisieren". Für die $k=4$ -Profile sind die Unterschiede im Vergleich zu ihrer Steilheit vernachlässigbar. Für die $k=2$ -Profile schien doch ein Test angebracht:

Als Punktwert für jede Vp wurde die Summe der Wiederholungen nach r und Nichtwiederholungen nach f, also der in der Hauptinstruktion definierte Lernerfolg, berechnet. Als mittlere Lernerfolge ergaben sich für die akt-2-Gruppe 30,9 bzw. 64,4% und für die nver-2-Gruppe 34,0 bzw. 70,8%. (Die Prozentangaben resultieren aus der Division durch 48: 12r- und 12f-Items, 2 Listen.) Der U-Test ergab Nichtsignifikanz dieses Unterschiedes (s. SACHS, 1969, S. 293; $\hat{U}=65,5 > U(14;13;5\%)=50$). Es wurde zweiseitig getestet, da a priori keine einseitige Fragestellung begründet war (s. S. 167). Wegen der erhaltenen Nichtsignifikanz entfällt die Problematik eines erst anhand der Daten beschlossenen Tests.

Erinnerungsdaten

In der unteren Hälfte der Abb. 7.1.3-1 sind graphische Zusammenfassungen der Erinnerungsdaten dargestellt. Die dritte Reihe, (H) bis (J), zeigt die wichtigsten Ergebnisse dieses Experiments. Über den Rückmeldungen r und f sind die relativen Effizienzmaße $re(r!)$ und $re(f!)=1-re(r!)$ aufgetragen. Natürlich ist die Angabe von $re(f!)$ vollkommen redundant, diese Darstellung wurde aber der Anschaulichkeit halber und wegen der Vergleichbarkeit mit den darunter dargestellten Graphiken (K) bis (M) gewählt. Auf der Abzissenachse ist die Rückmeldung b weggelassen, da b für $re(r!)$ ja ausdrücklich keine Rolle spielt.

Die Bilder (H) und (I) sind für die Fragestellung (5a) bzw. die zugehörigen vier Hypothesen relevant (s. S. 164f.). Hypothesisch vorhergesagte Symmetrie, d.h. $re(r!)=50\%$, zeigt sich durch ein bis auf Zufallsschwankungen waagerechtes Profil an, Asymmetrie zugunsten von r dagegen durch ein von r nach f abfallendes Profil. Dem Augenschein nach sind die vier Hypothesen gut bestätigt. - Wie auf S. 165 (Punkt 1) gesagt wurde, können die Profile der $k=2$ -Gruppen als eine Art Norm dienen, mit denen eine etwaige Symmetrie in der beob-4-Gruppe zu vergleichen wäre. Diese Möglichkeit braucht allerdings nicht besonders in Anspruch genommen zu werden, da das beob-4-Profil nicht einmal der Tendenz nach eine Asymmetrie zugunsten von r erkennen läßt.

Die augenscheinliche Bestätigung der Hypothesen wurde statistisch abgesichert (WILCOXON-Test; s. SACHS, 1969, S. 312). Für die akt-4-Gruppe ergab sich: $\hat{R}=14 < R(19;5\%)_{\text{eins.}}=53$, d.h. signifikante Abweichung von $re(r!)=50\%$. Für die drei anderen Gruppen zeigte sich keine bedeutsame Abweichung von $re(r!)=50\%$: akt-2: $\hat{R}=29 > R(11;5\%)_{\text{zweis.}}=10$; beob-2: $\hat{R}=26,5 > R(11;5\%)_{\text{zweis.}}=10$; beob-4: $\hat{R}=25,5 > R(11;5\%)_{\text{zweis.}}=10$.

Zu diesen drei letzten Gruppen ist zu bemerken: Bekanntlich kann man Nullhypothesen nicht verifizieren, die numerischen Verhältnisse aber liegen sehr günstig für die Hypothese $re(r!) = 50\%$. Über die akt-2-, beob-2- und beob-4-Gruppen

gemittelt ergibt sich: $re(r!) = 49,3\%$. (Wird das nreh-2-Ergebnis aus Bild (J) hinzugenommen, ist der Unterschied zu 50% noch geringer: $re(r!) = 49,8\%$.)

Die unter den Bildern (H) und (I) dargestellten Graphiken (K) und (L) zeigen im wesentlichen dieselben Verhältnisse: Nur bei den akt-4-Profilen, sowohl den R-recall- als auch O-recall-Profilen, zeigt sich eine deutliche Abweichung von der Waagerechten, und zwar im Sinne einer Asymmetrie zugunsten von r.

Diese R- und O-recall-Daten sind, wie schon gesagt, nur deshalb zusätzlich zu $re(r!)$ dargestellt worden, um die Vergleichbarkeit mit den Experimenten I und II herzustellen. In den Abbildungen auf S. 106 und 108 sind auch die bb' - und bb'' -Werte aufgetragen, die für den augenblicklichen Vergleich aber keine Rolle spielen. Die zu den Bildern (K) bis (M) gehörenden bb' - und bb'' -Daten werden daher auch nicht mehr hier angegeben (wie nach der Hauptinstruktion, die ja die Aufmerksamkeit auf r und f lenkte, zu erwarten ist, liegen diese Häufigkeiten beträchtlich tiefer als die Erinnerungshäufigkeiten bezüglich der r- und f-Items). - In der anschließenden Diskussion (7.1.4) braucht nicht mehr gesondert auf die R- und O-recall-Daten eingegangen zu werden.

Es sind jetzt noch die zu der nreh-Bedingung gehörenden Bilder (J) und (M) zu besprechen. (Die punktierten Profile, "nnreh-4"-Bedingung, beziehen sich auf das später darzustellende Experiment IV.)

Zu Bild (J): Der WILCOXON-Test ergab eine signifikante Abweichung von $re(r!)=50\%$ für die nreh-4-Gruppe ($\hat{R}=2 < R(9;5\%)$ eins. =8) und keine Signifikanz für die nreh-2-Gruppe ($\hat{R}=22 > R(10;5\%)$ zweis. =8). Die Verhältnisse liegen sehr ähnlich wie bei den akt-Profilen in Bild (H); insbesondere sind unter der theoretisch interessanteren k=4-Bedingung die Profile, gegen die Erwartung, praktisch gleich (s. S. 167): $re(r!)=70,4\%$ (akt) vs. $re(r!)=69,1\%$ (nreh).

Ebenso ist auch das Bild (M) dem Bild (K) recht ähnlich, jedenfalls was die (A-)Symmetrieaussagen betrifft. Auffällig

sind allerdings die stark erhöhten Werte $rr'!$ und $ff'!$ unter $k=2$ (man betrachte dazu noch einmal das unerwartet steile $nreh-2$ -Profil aus Bild (F)).

Die Prozentangaben in der oberen Hälfte der Tab. 7.1.3-1 geben an, wie relativ häufig r bzw. f im Sinne der Definition auf S. 157 *w i r k s a m* geworden sind. Dabei ist z.B. $n'(R-R!r!) = n(R-R!r!)/N \cdot 24$; $n(R-R!r!)$ ist die absolute Häufigkeit des Auftretens von $(R-R!r!)$ innerhalb einer Versuchsgruppe, N ist die Gruppengröße, 24 ergibt sich aus den je 12 r -Items der Listen A' und B'.

Die Größe $n'(. - R!b!)$ ist analog gebildet. $n'(. - R!b!)$ setzt sich dabei aus $n(R-R!b!)$ und $n(\bar{R}-R!b!)$ zusammen; diese unterschiedlichen Fälle werden im Zusammenhang mit Tab. 7.1.3-3 diskutiert. $n'(. - R!b!)$ ist nur der Vollständigkeit halber angeführt, hat allerdings auch eine ergänzende illustrative Funktion.

Aus den Größen $n'(R-R!r!)$ und $n'(\bar{R}-R!f!)$ lassen sich im übrigen die relativen Effizienzmaße $re(r!)$ der Bilder (H) bis (J) in Abb. 7.1.3-1 berechnen ($N \cdot 24$ kürzt sich in der Definitionsformel auf S. 158 heraus); z.B. für die akt-2-Gruppe: $re(r!) = 25,6 / (25,6 + 27,1) = 48,6 (\%)$.

Die Angabe dieser Größen ist trotzdem nicht redundant. Will man zwei Gruppen unmittelbar bezüglich der relativen Effizienz von r und f vergleichen, wie es der Fragestellung (5a) entspricht, dann ist $re(r!)$ vorzuziehen. Interessiert man sich aber dafür, in welchem Ausmaß r bzw. f innerhalb einer Gruppe *w i r k s a m* geworden sind, so sind $n'(R-R!r!)$ bzw. $n'(\bar{R}-R!f!)$ für die Beschreibung günstiger.

Man könnte auch, statt den Lernerfolg (LE) durch die Steilheit der Profile aus den Wiederholungsdaten zu charakterisieren, z.B. folgendes Lernerfolgsmaß definieren:

$$LE' = n'(R-R!r!) + n'(\bar{R}-R!f!).$$

Dieses hat den Vorteil, daß es sich besser als LE in den Rahmen der anderen Größen wie $re(r!)$ und $\emptyset$ einpaßt (s. dazu die Anwendungen auf S. 215 und 223).

Tab. 7.1.3-1: Relative Häufigkeiten $n'(\dots)$ (in%) der Fälle $(R-R!r!)$, $(R-R!f!)$ und $(-R!b!)$.
Interspezifische Assoziationskoeffizienten ϕ .
Experiment III; "nnreh" aus Experiment IV; weitere Erläuterung s. Text.

		akt	beob	nreh	nnreh
k=2	$n'(R-R!r!)$	25,6	17,8	29,6	
	$n'(R-R!f!)$	27,1	16,1	27,6	/
	$n'(-R!b!)$	11,9	8,9	13,6	
k=4	$n'(R-R!r!)$	28,4	13,3	23,3	9,4
	$n'(R-R!f!)$	11,9	15,3	10,4	4,9
	$n'(-R!b!)$	2,7	6,9	3,8	0,7
k=2	ϕ_r	0,79	0,71	0,73	
	ϕ_f	0,83	0,63	0,83	/
	ϕ_b	0,40	0,39	0,51	
k=4	ϕ_r	0,76	0,58	0,77	0,43
	ϕ_f	0,60	0,70	0,61	0,36
	ϕ_b	0,28	0,58	0,41	0,02

Tab. 7.1.3-2: Zum Vergleich der relativen Effizienzmaße (in %) $re(b)$ (s.S.44) und $re(r!)$ (s.S.158).
Experiment III; "nnreh" aus Experiment IV.

Position	akt		beob		nreh		nnreh	
	$re(b)$	$re(r!)$	$re(b)$	$re(r!)$	$re(b)$	$re(r!)$	$re(b)$	$re(r!)$
k=2	I	40,0	50,0	55,2	57,6	36,1	47,1	
	II	29,8	47,2	51,4	47,6	10,4	55,4	/
	I+II	34,0	48,6	53,2	52,4	24,8	51,6	
k=4	I	82,2	69,2	90,6	51,2	14,8	69,8	60,0
	II	85,1	71,6	51,2	43,4	48,6	68,6	34,8
	I+II	83,8	70,4	69,0	46,6	33,9	69,1	39,3

Tab. 7.1.3-3: Schätzungen (in %) der Beharrungsparameter g' (bei Erinnerung an b) und g'' (bei Vergessen von 0).
Experiment III; "nnreh" aus Experiment IV.

		akt	beob	nreh	nnreh
k=2	g'	75,8	84,2	74,5	
	g''	80,9	68,1	54,0	/
k=4	g'	37,0	53,0	40,8	40,0
	g''	57,7	63,6	32,4	27,0

In der unteren Hälfte der Tab. 7.1.3-1 sind die interspezifischen Assoziationskoeffizienten ϕ nach COLE eingetragen (s. S. 185f.). Erwartungsgemäß sind alle Koeffizienten positiv und sogar von beträchtlicher Höhe. Daß ϕ_b vergleichsweise klein ausfällt, war ebenfalls zu erwarten. (Die nnreh-Bedingung soll noch außer Betracht bleiben.)

Unerwartet dagegen ist, daß die Koeffizienten der nreh-Gruppen noch nicht einmal der Tendenz nach niedriger als die der akt-Gruppen liegen (s. S. 168). Sie sind vielmehr erstaunlich ähnlich.

Wegen dieser Ähnlichkeit (s. auch die Bilder (H) und (J) in Abb. 7.1.3-1) sollen die nreh-Gruppen in den folgenden Vergleich der (A-)Symmetriegruppen mit eingeschlossen werden. Die Vorhersage (s. S. 168): $\phi_r = \phi_f$ bei Symmetrie, wird von der akt-2-Gruppe, und die Vorhersage: $\phi_r > \phi_f$ bei Asymmetrie, wird von den akt-4- und nreh-4-Gruppen hinreichend unterstützt. Bei den anderen Bedingungen sind die Verhältnisse nicht ganz so klar, es ergibt sich allerdings auch kein starker Trend gegen die Erwartung. Mittelt man über alle Symmetriegruppen (akt-2, beob-2, beob-4, nreh-2) und Asymmetriegruppen (akt-4, nreh-4), so erhält man: bei Symmetrie $\phi_r = 0,70$ und $\phi_f = 0,75$; bei Asymmetrie $\phi_r = 0,77$ und $\phi_f = 0,61$. Das Resultat ist in Anbetracht der zu erwartenden Zufallsschwankungen (s. S. 169) einigermaßen befriedigend.

In Tab. 7.1.3-2 sind die beiden relativen Effizienzmaße $re(b)$ und $re(r!)$ zusammengestellt (für die nver-Gruppen läßt sich $re(r!)$ nicht berechnen; sie sind daher in der Tabelle nicht aufgeführt). Die $re(b)$ -Maße sind für die Positionen I und II aus den Profilen der Bilder (A) bis (C) in Abb. 7.1.3-1 berechenbar, für die zusammengefaßten Positionen I+II aus den Profilen der Bilder (E) und (F). Dieser Hinweis dient nur der Veranschaulichung von $re(b)$. Ebenso sind die $re(r!)$ -Maße für I+II in den Bildern (H) bis (J) graphisch dargestellt.

Die Tabelle dient dem Vergleich der Nützlichkeit beider Effizienzmaße. Konzentriert man sich zunächst auf die Zusammenfassung der Positionen I+II, so schwankt $re(r!)$ in den Symme-

triegruppen (akt-2, beob-2, beob-4, nreh-2) nur geringfügig um den Erwartungswert von 50% (von 46,6% bis 52,4%), während re(b) in diesen Gruppen von 24,8% bis 69,0% variiert. Für die Asymmetriegruppen (akt-4, nreh-4) schwankt re(r!) nur zwischen 69,1% und 70,4%, re(b) dagegen zwischen 33,9% und 83,8%. (Bei Hinzunahme der nnreh-Bedingung ist das Ergebnis nur geringfügig schlechter für re(r!).)

Die eben behandelte Gegenüberstellung von re(b) und re(r!) wurde zwischen Gruppen durchgeführt, die sich durch ihre experimentellen Bedingungen unterscheiden. Es soll nun noch ein Schwankungsvergleich innerhalb der Gruppen vorgenommen werden. Dazu werden die Differenzen von re(b) zwischen den Positionen I und II gebildet und über alle Gruppen (einschließlich nnreh) gemittelt. Es ergab sich als mittlere Variationsbreite für re(b): 20,2%. Dasselbe Verfahren führte bei re(r!) zu nur 6,4%.

In Tab. 7.1.3-3 sind die Schätzungen für die Parameter g' und g'' enthalten. g' (bzw. g'') ist ein Maß für die Beharrung auf einer Antwort R bei Erinnerung an b (bzw. bei Nichterinnerung an 0). Da die Prozentzahlen etwas unübersichtlich sind, soll über die akt-, beob- und nreh-Gruppen gemittelt werden. Daß etwaige Wechselwirkungen, z.B. die trendmäßig höheren Werte unter der beob-Bedingung, dabei unter den Tisch fallen könnten, soll zumindest angemerkt werden. Als gewichtete Mittelwerte ergeben sich (in %):

	g'	g''
k=2	78,2	63,8
k=4	42,8	54,6

Die Beharrungsparameter g' und g'' über k=2,4 gemittelt unterscheiden sich nicht deutlich. Die Schätzungen für k=2 liegen dagegen beträchtlich höher als die für k=4; hier scheint eine Aufspaltung in einen Zufallsanteil $1/k$ und einen systematischen "wahren" Beharrungsanteil g_w angebracht (s. dazu S. 139f.), also $g' = 1/k + g'_w$ bzw. $g'' = 1/k + g''_w$.

	g'_w	g''_w	
k=2	28,2	13,8	(21,0)
k=4	17,8	29,6	(23,2)
	(23,0)	(21,6)	

In den runden Klammern (...) stehen die gemittelten Randsummen.

Nachdem die Zufallsanteile $1/k$ abgespalten wurden, sind die systematischen "wahren" Komponenten für $k=2$ und $k=4$ von derselben Größenordnung. Die vier Zellenwerte lassen eine gewisse Wechselwirkung zwischen $k=2,4$ und g'_w , g''_w erkennen, für die sich aber nur sehr spekulative Interpretationen anbieten. Die gemittelten Randsummen unterscheiden sich so wenig, daß es vernünftig erscheint, einen einzigen Mittelwert zu berechnen: $g_w = 22,3\%$. - g_w ist also der allgemeine systematische Behandlungsanteil, unabhängig von k und ' bzw. ".

Man beachte die vielen Mittelwertbildungen, die von den sechs Gruppen der Tab. 7.1.3-3 mit insgesamt 87 Vpn und je zwei Listen zu diesem einzigen Kennwert führten. Auf die dadurch möglicherweise verdeckten Wechselwirkungen wurde schon hingewiesen.

Das Ergebnis dieses relativ groben Verfahrens ist also, daß kein Unterschied zwischen g' und g'' gemacht zu werden braucht, d.h. eine Unterstützung von BUCHWALDs diesbezüglicher Annahme (s. S. 30f.); es sei aber an die Instruktionsabhängigkeit der Bedeutungsauffassung von b durch die Vpn erinnert, die sich prinzipiell auch auf den Wert von g' übertragen kann. - Ferner scheint eine Aufspaltung von g in $1/k$ und g_w sinnvoll zu sein, also $g = 1/k + g_w$. Für $k=2$ bzw. 4 und mit dem einheitlichen Wert $g_w=22,3\%$ gerechnet ergibt sich damit: $g_{k=2}=50\%+22,3\%=72,3\%$ und $g_{k=4}=25\%+22,3\%=47,3\%$.

(Man vergleiche damit die auf einem ganz anderen Weg und mit anderen Vpn gewonnenen Schätzwerte von g auf S. 109.)

Entsprechend der quasi-quantitativen Betrachtung auf S. 139f. ergibt sich im Rahmen des BM:

$re(b, k=2)=100\%-72,3\%=27,7\%$ und $re(b, k=4)=100\%-47,3\%=52,7\%$

in qualitativer Übereinstimmung mit den dortigen Überlegungen.

Es sei jedoch einschränkend betont, daß die Ableitung der Beziehung $re(b)=1-g$ auf S. 44 das strenge BM voraussetzte, daß aber den hier betrachteten Gruppen Rehearsal erlaubt war. (Es fand wohl auch unter der nreh-Bedingung beträchtliches Rehearsal statt.)

Die für die Rückmeldungen r, f und b ausgewerteten a-posteriori-Wahrscheinlichkeiten (s. S. 187f.) $p(\text{korrekte Erinnerung/gegebenes Konfidenzurteil})$, kurz $p(\text{korrr./Konf.})$, sind in Abb. 7.1.3-2 als Profile dargestellt. Für das theoretische Zufallsniveau wurde 33,3% (= 1/3) angenommen; die Werte über den Kategorien "unsicher" und "sicher" sind empirisch gefunden.

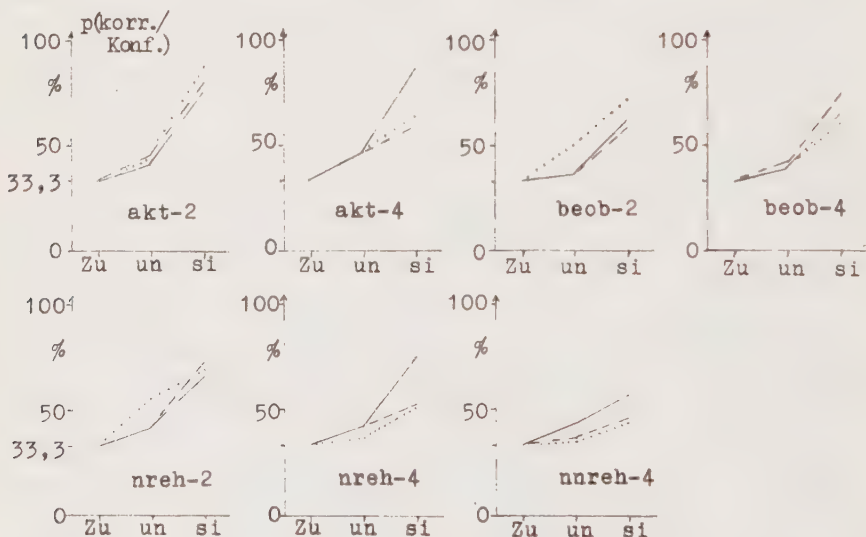


Abb. 7.1.3-2: a-posteriori-Wahrscheinlichkeiten $p(\text{korrr./Konf.})$ für die Rückmeldungen r: —, f: ---- und b:
Konfidenzstufen: "Zufall" (Zu), "unsicher" (un) "sicher" (si). (Experiment III; "nnreh-4" aus Experiment IV)

Die Profile verhalten sich ordnungsgemäß, d.h. sie steigen mit zunehmender subjektiver Sicherheit an. Daraus folgt, daß

die Vpn die Sicherheitskategorien adäquat benutzt haben. Die Konfidenzstufe "sicher" hebt sich deutlich gegenüber den beiden anderen Kategorien ab. Dieses Ergebnis ist eine nachträgliche Rechtfertigung für die bevorzugte Verwendung der "sicheren" Daten, wie z.B. in $re(r!)$ bzw. bei der Ableitung der Latenzzeithypothesen (s. S. 157f.). - Das allgemeine Niveau ist bei den akt-Gruppen etwas höher als bei den übrigen. Das kann daran liegen, daß die akt-Bedingung für die Vpn wohl die unkomplizierteste ist, d.h. ohne Zeitdruck oder zusätzliche beob-Instruktion.

Auffällig und unerwartet ist, daß sich die Symmetriegruppen (akt-2, beob-2, beob-4, nreh-2) deutlich von den Asymmetriegruppen (akt-4, nreh-4) unterscheiden. Während die r-, f- und b-Profile allgemein relativ nahe beieinander liegen, trennen sich unter akt-4- bzw. nreh-4-Bedingungen die r- von den f- und b-Profilen. In den Asymmetriegruppen wurde also das Konfidenzurteil "sicher" bei r-Items adäquater angegeben als bei den f- und b-Items; oder anders ausgedrückt: es wurde von den Vpn relativ häufig z.B. "f!" behauptet, obwohl in Wirklichkeit r- bzw. b-Items vorlagen.

Latenzzeiten

Die Ergebnisse der Latenzzeitmessung sind in Tab. 7.1.3-4 enthalten. Die Reihenfolge der Mediane ist dabei so gewählt, daß sie der durchschnittlichen Rangreihenfolge unter der $k=2$ -Bedingung entspricht. Um den Vergleich und die Beurteilung der Latenzzeitmuster hinsichtlich der Hypothesen H_1 bis H_{10} zu erleichtern, sind die Mediane in Abb. 7.1.3-3 graphisch dargestellt.

In die Tabelle und die Abbildung wurden auch die Ergebnisse der bisherigen Experimente I und II und die der noch zu referierenden Experimente IV und V aufgenommen; diese Zusammenfassung ist für die unter 7.3 erfolgende Gesamtdiskussion nützlich. Bei der Reihenfolge wurde so vorgegangen, daß später zu vergleichende Bedingungen auch optisch möglichst nahe zusammenliegen.

In den angegebenen Zeitmedianen sind beide Übungsstadien (Listen A' und B') zusammengefaßt. Die vorherige Inspektion der Latenzzeitmuster hatte keinen irgendwie interpretierbaren Trend zwischen den Positionen I und II erkennen lassen.

Tab. 7.1.3-4: Gesamtübersicht über alle Experimente (I-V): Zeitmediane (in sec); T_k ist der Gesamtmedian aus T_{rr} .. $T_{b\bar{b}}$.

Exp.	k	Gruppe	T_{rr}	$T_{f\bar{f}}$	T_{bb}	T_{ff}	$T_{r\bar{r}}$	$T_{b\bar{b}}$	T_k
III	2	akt-2	4,51	4,77	5,44	5,54	5,93	5,49	5,47
	4	akt-4	4,48	6,01	5,46	5,97	6,18	5,98	5,98
	2	beob-2	4,13	4,48	4,73	5,60	5,10	5,73	4,92
	4	beob-4	4,78	6,02	5,36	5,60	6,33	6,22	5,81
	2	nver-2	2,54	2,66	2,98	3,06	3,38	3,36	3,02
	4	nver-4	2,68	3,05	2,80	3,15	3,34	3,78	3,10
	2	nreh-2	3,93	3,94	4,64	4,60	4,70	5,30	4,62
	4	nreh-4	4,67	4,80	4,79	4,88	4,53	5,23	4,80
I	2	R-corr(I)	2,22	2,30	2,45	2,53	2,44	2,55	2,45
	4	R-corr(I)	2,05	2,45	2,55	2,65	2,63	2,54	2,55
II	4	R-corr(A)	2,31	3,37	3,40	3,31	3,35	3,45	3,36
		R-corr(B; speed)	1,71	2,43	2,24	2,13	2,51	2,29	2,27
IV	4	nnreh-4	4,23	4,52	3,97	4,45	4,37	4,33	4,35
V	2	S-R-f	2,15	2,16	2,05	2,48	2,95	2,46	2,31
		S-R	2,04	2,23	2,34	2,35	2,50	2,57	2,35

In Tab. 7.1.3-5 ist die Beurteilung der Latenzzeitergebnisse hinsichtlich der Hypothesen H1-H10 in einem Überblick dargestellt. (Von den aufgeführten Experimenten wird zunächst nur das Experiment III betrachtet.)

Unter der Hypothesennummer steht der Inhalt der Hypothese in Kurzform. In der Spalte unter H1 ist durch <bzw.> die jeweils gefundene Größer-Kleiner-Beziehung zwischen T_{ff} und $T_{r\bar{r}}$ (in dieser Reihenfolge) gekennzeichnet; es ist unwahrscheinlich, auf der kontinuierlichen Zeitskala Gleichheit zwischen den beiden Medianen zu finden. Die übrigen Spalten enthalten

Exp. k Gruppe

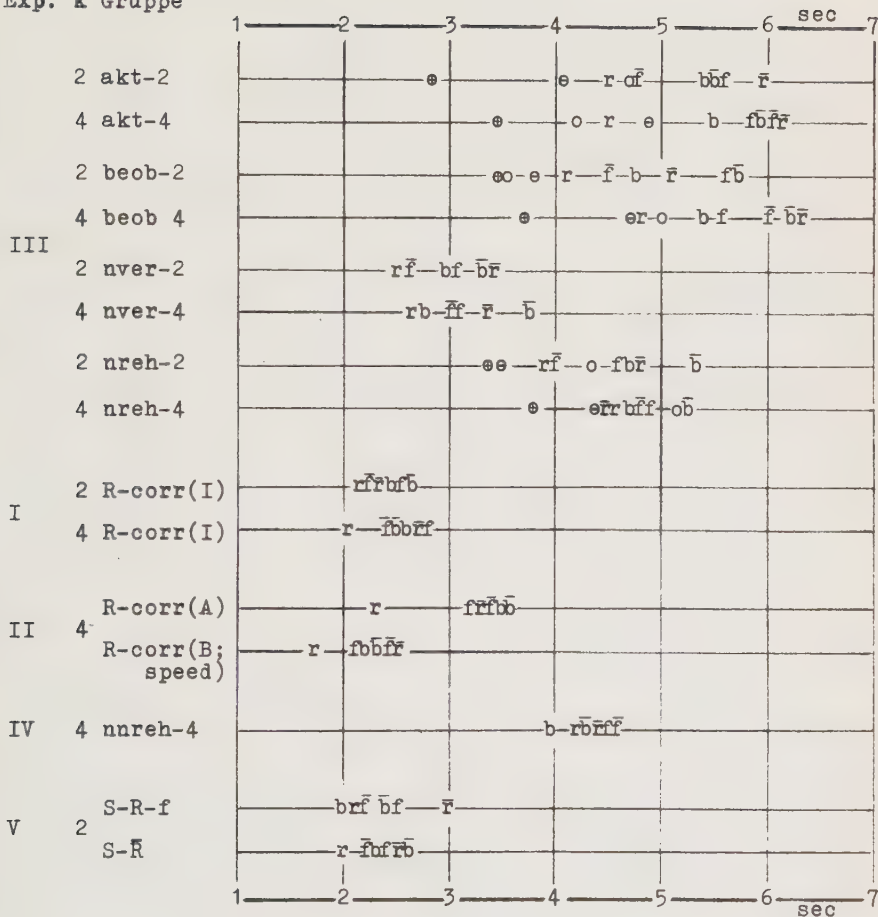


Abb. 7.1.3-3: Die Mediane T_{rr} , $T_{r\bar{r}}$, T_{ff} , $T_{f\bar{f}}$, T_{bb} , $T_{b\bar{b}}$ sind der Kürze halber nur durch die sie unterscheidenden Buchstaben dargestellt, also r, $\bar{r}$, f, $\bar{f}$, b, $\bar{b}$. Die Symbole e, e und o beziehen sich auf die "sicheren" (c,1-u)-Wege; weitere Erklärung im Text auf S. 205f.

ein "+", wenn der empirische Befund der Hypothese entspricht, sonst ein "-".

Die für die Beurteilung der Hypothesen H7 bis H10 benötigten Differenzen wurden aus Tab. 7.1.3-4 berechnet und in die entsprechenden Zeilen und Spalten eingetragen.

Die Resultate sprechen dem Augenschein nach für den Hypothesenkomplex H2 bis H9 (H1 und H10 sollen im Augenblick ausgeklammert bleiben).

Selbstverständlich ist das Verhältnis der "+" zu den "-" nur ein grober Indikator; es müßte z.B. die partielle Abhängigkeit der Hypothesen untereinander berücksichtigt werden, sowohl was ihre theoretische Ableitung anbetrifft als auch hinsichtlich der Tatsache, daß die Mediane innerhalb der einzelnen Gruppen von denselben Vpn stammen und zur Beurteilung der Hypothesen mehrfach herangezogen werden. Ferner kann ein einziger "Ausreißer" das Gesamtbild stören, wie z.B. in der nreh-4-Gruppe der Median $T_{r\bar{r}}$, der, verglichen mit den anderen Gruppen, sehr auffallend an erster Stelle liegt (s. Abb. 7.1.3-3).

Die meisten negativen Befunde findet man in den nreh-Gruppen und zu der Hypothese H8.

Bei H1 handelt es sich um eine "zu bestätigende Nullhypothese". Das Ergebnis der Gesamtauswertung (s. 7.3) sei vorweggenommen: Nach dem WILCOXON-Test (s. S. 232) ergab sich keine Signifikanz für das überzufällige Auftreten einer der beiden Beziehungen "<" oder ">". Die Hypothese $H1: T_{ff} = T_{r\bar{r}}$ kann also beibehalten werden. Es ist zu erinnern, daß sowohl T_{ff} als auch $T_{r\bar{r}}$ Schätzungen des GK-Medians sein sollen (s. S. 71).

Ergebnisse zum "Zeitlupeneffekt" (s. S. 160f.): Dieser vorhergesagte Effekt bezog sich auf das Verhalten der Latenzzeitmuster beim Wechsel vom Nichtverbalisieren zum Verbalisieren im Prüfdurchgang. Vergleicht man in Abb. 7.1.3-3 die nver-2- mit der akt-2-Gruppe und die nver-4- mit der akt-4-Gruppe, so stellt man erwartungsgemäß eine Verschiebung zu

Tab. 7.1.3-5: Beurteilung der Latenzzeitergebnisse hinsichtlich der Hypothesen H1 bis H10.
Die Mediane T_{rr} , $T_{\bar{r}}$, T_{ff} , T_{bb} und $T_{b\bar{b}}$ sind der Kürze halber nur durch die sie unterscheidenden Buchstaben dargestellt, also durch r, $\bar{r}$, f, $\bar{f}$, b und $\bar{b}$. Weitere Erläuterung im Text auf S.200 und in 7.3

Exp. k Gruppe	H1 $f=\bar{r}$	H2 $r < \bar{r}$	H2' $r < f$	H3 $r < \bar{f}$	H4 $b < \bar{r}$	H4' $b < f$	H5 $r < b$	H6 $b < \bar{b}$	H7 $\left(\frac{\bar{f}-r}{f-r}\right)^2 < \frac{b}{4}$	H8 $\left(\frac{b-b}{b-b}\right)^2 < \frac{T_2}{T_4}$	H9 $T_2 < \frac{T_2}{T_4}$	H10 $\left(\frac{\bar{f}-r}{f-r}\right)_{\text{asym}} < \left(\frac{\bar{f}-r}{f-r}\right)_{\text{asym}}$
2 akt-2	<	+	+	+	+	+	+	+	0,26 +	0,05 +	+	1,53 +
4 akt-4	<	+	+	+	+	+	+	+	1,53	0,52	+	
2 beob-2	>	+	+	+	+	+	+	+	0,35 +	1,00 -	+	
4 beob-4	<	+	+	+	+	+	+	+	1,24	0,86	-	
2 nver-2	<	+	+	+	+	+	+	+	0,12 +	0,38 +	+	
4 nver-4	<	+	+	+	+	+	+	+	0,37	0,98	+	
2 nreh-2	<	+	+	+	+	-	+	+	0,01 +	0,66 -	+	
4 nreh-4	>	-	+	+	-	+	+	+	0,13	0,44	-	
I 2 R-corr(I)	>	+	+	+	-	+	+	+	0,08 +	0,10 -	+	
4 R-corr(I)	>	+	+	+	+	+	+	-	0,40	-0,01	-	
II 4 R-corr(A)	<	+	+	+	-	-	+	+	5	2	5	:"Treffer"x
4 R-corr(B; speed)	<	+	+	+	+	-	+	+	3,13	81,25	3,13	
IV 4 nreh-4	>	+	+	+	+	+	-	+				
V 2 S-R-f	<	+	+	+	+	+	-	+				
S-R	<	+	+	+	+	+	+	+				
*	*	14	15	15	12	12	13	14	:"Treffer"x			

$P(X \geq x/n=5; p=0,5)$
(in %)

(vgl. 7.3)

*WILCOXON-Test: nichtsignifikant (s.S.232).

längeren Zeiten und eine Auffächerung des Zeitmusters fest. Die Rangreihenfolge der Mediane bleibt dabei mehr oder weniger gut erhalten. Als deskriptives Maß für die Ähnlichkeit der Rangreihen zwischen den nver- und akt-Gruppen wurde der SPEARMANsche Rangkorrelationskoeffizient r_S berechnet. Es ergab sich:

$$r_S(\text{nver-akt-2})=0,94 \text{ und } r_S(\text{nver-akt-4})=0,71.$$

Als Mindestanforderung für die Stützung des Zeitlupeneffekts war die gleich gute Unterstützung der Hypothesen H1 bis H9 in den nver- und akt-Gruppen genannt worden (s. S. 170). Diese Mindestanforderung wird vollständig erfüllt (s. Tab. 7.1.3-5).

Da die Absicherung des Zeitlupeneffekts für die mögliche Übertragbarkeit der unter Verbalisierungsbedingungen erhaltenen Befunde auf Nichtverbalisierungssituationen wichtig ist, sollen zumindest beispielhaft die Latenzzeitverteilungen unter beiden Bedingungen noch miteinander verglichen werden. Es genügt vorläufig, wenn die Form der Verteilungen nicht allzu kraß voneinander abweicht; für diesen Zweck können auch die kritischen Anmerkungen zur partiellen "inneren" Abhängigkeit der Häufigkeitspolygone vernachlässigt werden (s. S. 99f.). Als Beispiele wurden die zum Median T_{rr} gehörenden Verteilungen gewählt (s. Abb. 7.1.3-4). Um die Form der Verteilungen besser vergleichen zu können, sind Nullpunkt und Maßeinheit auf den Abzissenachsen so gewählt, daß die Polygone bezüglich Median und Ausdehnung einigermaßen übereinstimmen. (Daß sich die nver- bzw. akt-Verteilungen hinsichtlich zentraler Tendenz und Ausdehnung unterscheiden, braucht nicht noch erst gesondert dargestellt zu werden: Das ist in der Aussage des Zeitlupeneffekts enthalten.)

Die Ähnlichkeit der Polygone in Abb. 7.1.3-4 unter den nver-2- und akt-2-Bedingungen ist augenfällig. Unter der k=4-Bedingung deutet sich ein Unterschied hinsichtlich der Schiefe an. - Zusammen mit den Werten von r_S betrachtet, scheint der Zeitlupeneffekt unter k=2 reiner ausgeprägt als unter k=4. (Man vergleiche dazu Bild (G) in Abb. 7.1.3-1 auf S. 189:

Hier war, deskriptiv gesehen, unter der $k=4$ -Bedingung eine größere Ähnlichkeit zwischen den nver- und akt-Profilen gefunden worden.)

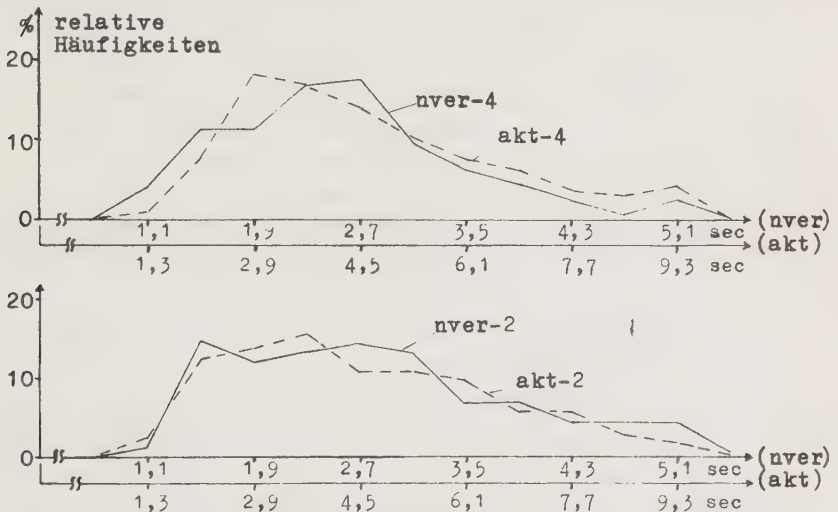


Abb. 7.1.3-4: Zum "Zeitlupeneffekt" zwischen den nver- und akt-Bedingungen: Vergleich der Zeitverteilungsformen (z.B. für rr-Antworten). Man beachte die unterschiedlichen Maßeinheiten auf den Abzissenachsen. (Experiment III)

Die Ableitung der Latenzzeithypothesen in Kapitel 4 (s. S. 70; s. auch S. 171) hatte sich besonders auf den Anteil "sicherer" (c,1-u)-Zeiten an der jeweiligen Gesamtverteilung bezogen. Es war speziell angenommen worden, daß diese Zeiten, trotz eventuellen Umschaltens wie z.B. in den ($\bar{R}$ -R!f!)-Fällen, relativ kurz seien gegenüber den längeren GK-Zeiten. Diese Annahme kann aufgrund der Verbalisierungsdaten überprüft werden.

Dazu wurden nur die Zeiten betrachtet, die zu den Fällen (R-R!r!) bzw. ($\bar{R}$ -R!f!) bzw. (R-R!b!) gehören (s. die Begründungen der Hypothesen H2, H3 und H4 auf S. 71f.). Die zugehörigen Mediane sind in Abb. 7.1.3-3 auf S. 201 eingetragen. Die Verschlüsselung ist folgende: ● für den (R-R!r!)-, ○ für

den $(\bar{R}-R!f!)-$, o für den $(R-R!b!)-$ Median.

Erwartungsgemäß sind diese "schnellen" Mediane, teilweise sogar beträchtlich, kleiner als die jeweils dazugehörenden Gesamtmediane T_{rr} , T_{ff} und T_{bb} . Die einzige Ausnahme findet sich in der nreh-4-Gruppe: Dort ist der $(R-R!b!)-$ Median größer als T_{bb} . Ebenfalls der Erwartung entsprechend ist die Reihenfolge: Der $(R-R!r!)-$ Median liegt jeweils an erster Stelle. Wegen der Ungewißheit über die Umschaltzeit in $(\bar{R}-R!f!)-$ Fällen war über das Verhältnis der $(\bar{R}-R!f!)-$ und $(R-R!b!)-$ Mediane zueinander keine Vermutung aufgestellt worden.

Es wurden auch noch die ebenfalls zum $(c_b, 1-u_b)$ -Weg gehörenden $(\bar{R}-R!b!)-$ Mediane (s. dazu H6 auf S. 73) berechnet: Wegen des Umschaltens waren diese ausnahmslos größer als die $(R-R!b!)-$ Mediane derselben Gruppe.

Der Einfluß der "schnellen" $(c, 1-u)$ -Zeiten auf die jeweiligen Gesamtmediane hängt außer von der Lage dieser Zeiten auf der Zeitachse von den prozentualen Anteilen an den Gesamtverteilungen ab. Das war der wesentliche Punkt bei der Begründung der Hypothese H10 auf S. 171 gewesen. Die Abb. 7.1.3-3 auf S. 201 soll daher noch durch Tab. 7.1.3-6 ergänzt werden, in der die prozentualen Anteile der $(R-R!r!)-$, $(\bar{R}-R!f!)-$ und $(R-R!b!)-$ Zeiten an den zugehörigen $(T_{rr}-, T_{ff}-$ und $T_{bb}-)$ Gesamtverteilungen angegeben sind.

Tab. 7.1.3-6: Prozentuale Anteile "schneller" $(c, 1-u)$ -Zeiten an den zugehörigen Gesamtverteilungen.- $n(rr)$ z.B. bedeutet die absolute Häufigkeit der rr -Antwortzeiten in der zu T_{rr} gehörigen Gesamtverteilung, usw. (Experiment III)

Gruppe	$\frac{n(R-R!r!)}{n(rr)}$	$\frac{n(\bar{R}-R!f!)}{n(ff)}$	$\frac{n(R-R!b!)}{n(bb)}$
akt-2	37,2	42,8	17,9
akt-4	<u>51,2</u>	<u>15,8</u>	3,8
beob-2	30,1	28,3	16,2
beob-4	<u>29,1</u>	<u>19,1</u>	11,7
nreh-2	39,4	41,6	18,8
nreh-4	48,2	13,7	2,8

(Die Komplemente zu 100% stellen den Anteil der jeweiligen GK dar. - Die unterstrichenen Prozentanteile in der Tabelle beziehen sich auf die Diskussion der Hypothese H10 und werden weiter unten besprochen.)

Die Eintragungen in Tab. 7.1.3-6 sind nicht zu verwechseln mit denen aus der oberen Hälfte von Tab. 7.1.3-1 auf S. 194; dort wurde sich auf die Gesamtzahl z.B. der f-Items bezogen, also auf $n(ff)+n(f\bar{f})=N \cdot 24$ (s.S. 193) statt wie hier nur auf $n(f\bar{f})$ usw.

Daraus folgt, daß die Anteile in der mittleren Spalte für die $k=4$ -Gruppen relativ klein ausfallen müssen, da $n(f\bar{f})$ einen hohen Rateanteil ($=3/4$) enthält und im Nenner steht. Aus diesen numerischen Verhältnissen folgt weiterhin, daß die Fälle ($\bar{R}-R!f!$), in denen also f wirksam geworden ist, bezüglich der zugehörigen Latenzzeitverteilung bzw. T_{ff} vermutlich weniger Effekt haben werden als z.B. die Fälle ($R-R!r!$) bezüglich T_{rr} .

Trotz der schon öfter angewerkten Problematik, die in den partiell abhängigen Latenzzeitverteilungen liegt, soll hier wenigstens eine Verteilung bildlich dargestellt werden, in der die Zusammensetzung der Gesamtverteilung aus schnellen ($c,1-u$)-Zeiten und langsameren GK-Zeiten besonders illustrativ demonstriert werden kann (s. Abb. 7.1.3-5).

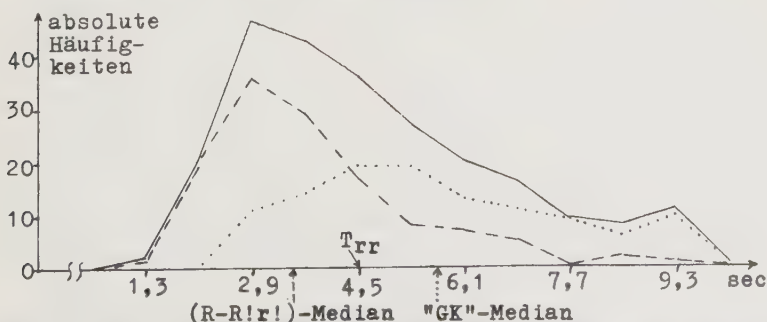


Abb. 7.1.3-5: Zusammensetzung der zu T_{rr} gehörigen Gesamtzeitverteilung (—) aus "schnellen" ($c,1-u$)-Zeiten (----) und "langsameren" GK-Zeiten (.....). (Experiment III; akt-4-Gruppe)

Es wurde die Verteilung zu T_{rr} der akt-4-Gruppe gewählt, da hier die beiden Anteile einigermaßen gleich sind (s. Tab. 7.1.3-6: 51,2% schnelle Zeiten vs. 48,8% langsamere GK-Zeiten).

Nachdem der Anteil der schnellen Zeiten aus der Gesamtverteilung herausgezogen ist, muß die (punktierte) Restverteilung die hypothetisch angenommene GK-Verteilung approximieren (s. S. 70). Andererseits sind auch die zu T_{rr} und T_{ff} gehörenden Verteilungen als Schätzungen von GK zu betrachten (s. H1, S. 71). Es seien nur die Mediane miteinander verglichen: (R-R!r!)-Median=3,46 sec; T_{rr} =4,48 sec; "GK"-Median=5,68 sec; T_{ff} =5,97 sec; T_{rr} =6,18 sec.

Der "GK"-Median erreicht zwar nicht ganz T_{ff} bzw. T_{rr} , liegt aber doch mehr in ihrer Nähe als bei T_{rr} . (Analoge Berechnungen für die anderen Versuchsgruppen zeigten teilweise bessere, teilweise schlechtere Übereinstimmungen mit der Erwartung.)

Es war auf S. 170 gesagt worden, daß Fragestellung (3), präzisiert durch H10 (s.S. 171), erst behandelt werden könnte, wenn die Fragestellungen (2) bzw. (5a) positiv beantwortet wären. Da die Fragestellung (5a) präzisierende Hypothesenschema auf S. 165 wurde zufriedenstellend durch die Daten unterstützt. Auch Fragestellung (2), speziell für die an der Formulierung von H10 beteiligten akt-4- und beob-4-Gruppen, wurde im wesentlichen positiv beantwortet, wenn auch noch nicht statistisch abgesichert (s. Tab. 7.1.3-5 auf S. 203).

Fragestellung (2) wird sogar noch über die aufgestellten Hypothesen H1 bis H6 hinaus durch die auffällig ähnliche Gesamtreihenfolge innerhalb der Latenzzeitmuster der akt-4- und beob-4-Bedingungen unterstützt. Als deskriptives Ähnlichkeitsmaß wurde der Rangkorrelationskoeffizient r_s nach SPEARMAN berechnet: $r_s(\text{akt-beob-4})=0,94$. Die Höhe dieses Koeffizienten weist auf den Latenzzeitmustern zugrundeliegende Gesetzmäßigkeiten hin, die durch die bisher aufgestellten Hypothesen H1 bis H6 noch nicht erfaßt wurden. (H7 bis H9 wurden nicht erwähnt, da es im Augenblick nur um die k=4-Bedingung geht.)

Trendmäßig wird H10 durch die beobachteten Daten unterstützt (s. Tab. 7.1.3-5 auf S. 203; s. auch Abb. 7.1.3-3). Der Effekt ist jedoch relativ gering, verglichen z.B. mit der Schwankung von T_{ff} zwischen den beiden Gruppen.

Betrachtet man in Abb. 7.1.3-3 (s. S. 201) für die akt-4- und beob-4-Gruppen die Lage der für die Begründung von H10 relevanten Mediane (σ und e) und in Tab. 7.1.3-6 (s. S. 206) die zugehörigen (unterstrichenen) Prozentanteile, so sind alle Voraussetzungen für eine Bestätigung von H10 gegeben. Offensichtlich aber haben die beim Übergang von akt-4 nach beob-4 sich verändernden Anteile an (R-R|r!)- und ($\bar{R}$ -R!f!)-Zeiten gegenüber der großen Varianz in den Gesamtzeitverteilungen zu wenig Gewicht, um einen starken Effekt hervorzurufen. (Zur Größenordnung der auftretenden Varianzen s. z.B. Abb. 7.1.3-4 auf S. 205)

Die weitere Inspektion der Latenzzeitmuster, die zum Hypothesenschema auf S. 165 gehören und prinzipielle Kontrollmöglichkeiten (s. S. 172) bieten könnten, erbrachte keine interpretierbaren Regelmäßigkeiten als Beitrag zur Beantwortung der Fragestellung (3), die nicht schon durch die Hypothesen H1 bis H10 erfaßt worden wären.

Die Ergebnisse des postexperimentellen Interviews (s. S. 180) werden für die Bedingungen akt-2, akt-4, beob-2, beob-4, nver-2 und nver-4 dargestellt. Dabei wurden die Gruppen von vornherein so zusammengefaßt, wie es dem Sinn der jeweiligen Frage entspricht (vor der Zusammenfassung war selbstverständlich auf etwaige Auffälligkeiten geachtet worden).

F1: 7% der beob-Vpn gaben an, r und f hätten emotional auf sie gewirkt; bei den akt-Vpn waren es 18%, bei den nver-Vpn 27%. Diese Reihenfolge ist nicht unerwartet; die akt-Vpn waren durch die Verbalisierungsanweisung mehr auf die Informationshaltigkeit der Rückmeldungen hingewiesen als die nver-Vpn. Im übrigen wurde die emotionale Wirkung sehr selten in einem direkten Sinne erlebt, sondern mehr in Hinblick auf z.B. günstig gewählte "Eselsbrücken": Das wurde von den Vpn betont.

F2: Auf 20% der Vpn wirkte die Rückmeldung b emotional positiv ("entlastend"), auf 31% negativ ("verunsichernd"), auf den Rest (49%) neutral (s. auch S. 19).

F3: Besonders auf r-Items konzentriert hatten sich 17% der k=2-Vpn, 27% der beob-4-Vpn und 50% der "aktiven" k=4-Vpn (akt-4, nver-4).

F4: b-Items zu lernen versucht, hatten 41% der Vpn aus den Verbalisierungsgruppen, aber nur 15% der nver-Vpn.

F5: Hinsichtlich allgemeiner Lernstrategien war, abgesehen von der Beschreibung individueller auf einzelne Items bezüglicher Merkhilfen, folgende Strategie im Lerndurchgang interessant, nämlich bei der Rückmeldung f sofort auf eine andere Alternative umzuschalten. Dieses Vorgehen wurde von k=4-Vpn praktisch nicht berichtet. Dagegen gaben in der nver-2-Gruppe 67% der Vpn diese Strategie an. In den Verbalisierungsgruppen (k=2) waren es nur 15%; weitere 26% aber hätten gerne diese Strategie benutzt, wenn ihnen nicht die Verbalisierungsinstruktion gegeben worden wäre: Wenn man bei der Sequenz "GOLD 5 f" versucht, sich ausschließlich "GOLD-3" einzuprägen, und dieses auch effektiv gelingt, kann man im Prüfungsdurchgang nur noch raten, ob im Lerndurchgang "GOLD 5 f" oder "GOLD 3 r" aufgetreten war.

Eine Vp aus der nver-2-Gruppe gab an, sich nur die Wörter gemerkt zu haben, bei denen "5" die richtige Antwort war, nachdem sie gegebenenfalls bei f umgeschaltet hatte. Bei allen anderen Wörtern, einschließlich b-Items, sagte sie im Prüfungsdurchgang dann "3". Nach dieser sehr effektiven Strategie brauchte die Vp sich im Lerndurchgang pro Lernliste durchschnittlich nur 12 Wörter zu merken (s. dazu auch HALL u. WENDEROTH, 1970, 1972; HERRMANN u. STAFF, 1973). Da dieses Vorgehen durch die Hauptinstruktion nicht verboten war, wurde diese Vp für die Auswertung beibehalten. Die nver-Bedingung sollte ja zur Kontrolle gegenüber den "Verbalisierungs"-Gruppen dienen, d.h. es mußten den nver-Vpn Strategien erlaubt sein, die andernfalls durch die Verbalisierungsanweisung unterdrückt wurden.

.1.4 Diskussion

Im großen und ganzen zeigten alle Profile aus den Wiederholungshäufigkeiten einen befriedigenden Verlauf, d.h. einen deutlichen Lernerfolg und monotonen Abfall von rr über bb nach ff.

Die sich abzeichnende Unterlegenheit der beob- gegenüber den akt-Gruppen braucht nicht diskutiert zu werden, da nicht die Abhängigkeit des L e r n e r f o l g s von der Lernmodalität in dieser Arbeit untersucht wird, sondern die Abhängigkeit der r e l a t i v e n E f f i z i e n z von r und f. Der niedrigere Lernerfolg der beob-4-Gruppe, relativ zur akt-4-Gruppe, resultiert besonders aus den Wiederholungshäufigkeiten rr; bei den ff-Daten zeigt sich dagegen eine, wenn auch sehr geringfügige, Überlegenheit der beob-4-Bedingung. Dieser vermutliche "Symmetrisierungs"-Effekt kann mit Hilfe der Erinnerungsdaten durchsichtiger gemacht werden (s. S. 214f.)

Auf die etwas unerwarteten Ergebnisse in den nreh-Gruppen wird im Zusammenhang mit der allgemeinen Diskussion der nreh-Bedingung eingegangen (s.u.).

Obwohl der Vergleich der nver-Profile mit den akt-Profilen keinen bedeutsamen Unterschied ergab, sollte, besonders für die k=2-Bedingung, doch der erhaltene höhere Lernerfolg im Auge behalten werden. Nach der Beantwortung der Frage F5 sind unter nver-Bedingungen andere, eventuell effektivere Strategien möglich als unter einer Verbalisierungsanweisung. Insbesondere unterläuft die erwähnte Strategie einer Vp, sich nämlich nur "5-Wörter" zu merken, vollkommen die Absicht des Experiments. Darauf soll in der allgemeinen Kritik des experimentellen Vorgehens noch einmal eingegangen werden (s. 8.4).

Nun zur Diskussion der Erinnerungsdaten: Die zur Fragestellung (5a) gehörenden Hypothesen (s. S. 164f.) wurden nachdrücklich unterstützt. Es spricht für die theoretische und praktische Nützlichkeit des vorgeschlagenen Effizienzmaßes $re(r!)$, daß in den Gruppen, in denen Symmetrie, d.h.

$re(r!) = 50\%$, vorausgesagt war, dieser theoretische Wert numerisch so gut und stabil angenähert wurde. (Zu den Symmetriegruppen ist auch die nreh-2-Bedingung zu rechnen, da auch für sie die Voraussage: $re(r!) = 50\%$ gilt.) Die in der akt-4-Gruppe gefundene Asymmetrie, $re(r!) = 70,4\%$, wird dadurch um so deutlicher.

Einschub: Dieses Ergebnis unter der akt-4-Bedingung wird weiter stabilisiert durch den fast gleichen Asymmetriewert, $re(r!) = 69,1\%$, der in der nreh-4-Gruppe gefunden wurde. - Im nachhinein scheint die Methode, Rehearsal in den nreh-Gruppen zu verhindern (s. S. 181), nicht glücklich gewählt zu sein: Die Vpn hatten, nach ihren Angaben, die ihnen zur Verfügung stehenden Zeiten durchaus zum Rehearsal genutzt. Danach kann man, mit Einschränkung, die nreh-4-Bedingung als eine Replikation der akt-4-Bedingung betrachten, jedenfalls was die Frage nach der relativen Effizienz von r und f anbetrifft, wie sie sich in den Werten von $re(r!)$ und ϕ_r, ϕ_f ausdrückt.

Auf jeden Fall kann der in 6.3.3 postulierte stark automatisierte Kontrollprozeß durch das Ergebnis unter der akt-4-Bedingung als gut unterstützt angesehen werden. Mögliche Alternativerklärungen, wie z.B. nach dem Effekt-Gesetz bzw. durch ein mögliches Mißverständnis der Instruktion und Lernaufgabe, sind ausgeschaltet: Die erstere durch den fehlenden Unterschied bezüglich $re(r!)$ zwischen den akt-2- und beob-2-Gruppen, die zweite durch die klare Symmetrie in der beob-4-Gruppe (s. dazu S. 165f.).

Die Interpretation der erhaltenen Symmetrie unter der beob-4-Bedingung ist weniger eindeutig. Die erste Interpretationsmöglichkeit ist die, daß bei "beobachtendem" Lernen der automatisierte Kontrollprozeß überhaupt nicht oder nur sehr schwach angestoßen wird, im letzteren Fall aber durch eine symmetriebetonende Instruktion überwunden werden kann. Dieses war der Standpunkt bei der Aufstellung der Hypothese $H(\text{beob-4})$ auf S. 151, der auch weiterhin favorisiert bleibt.

Als eine andere Erklärungsmöglichkeit muß jedoch im Auge be-

halten werden, daß auch unter der beob-4-Bedingung der automatisierte Kontrollprozeß ebenso angestoßen wurde wie unter der akt-4-Bedingung, daß aber in der beob-4-Gruppe die symmetriebetonende Instruktion wirksamer geworden ist (s. dazu S. 152). Ein schwacher Hinweis auf das eventuell stärkere "Durchschlagen" der Instruktion unter Beobachtungsbedingungen kommt von der Beantwortung der Frage F1 (S. 209): Auf beob-Vpn hatten demnach r und f am wenigsten emotional gewirkt, sie wurden vielleicht also weniger in der rationalen Informationsverarbeitung gestört.

Welchen Anteil auch immer die letztere Erklärungsmöglichkeit an dem symmetrischen Ergebnis unter der beob-4-Bedingung hat, es ist festzuhalten, daß beim Übergang von der akt-4- zur beob-4-Bedingung eine Symmetrisierung in der relativen Effizienz von r und f stattfand. Damit war das eine Ziel erreicht, zusammen mit den k=2-Gruppen eine günstige Ausgangsbasis für die Behandlung der Latenzzeiten, besonders der Fragestellung (3), zu konstruieren.

Wie wird nun durch die obigen Ergebnisse die Fragestellung (5b) bezüglich der Gültigkeit der Informationshypothese, z.B. nach BUCHWALD (1962), beantwortet?

Für die k=2-Bedingungen wird die rein kognitive Informationshypothese, d.h. Annahme gleicher Informationswerte für r und f, eindrucklich unterstützt. (Daß BUCHWALD, 1962, noch eine zusätzliche Strategie zugunsten von f annehmen muß, liegt an seiner Beurteilung der relativen Effizienz von r und f anhand des Wiederholungsniveaus nach b. Diese Zusatzstrategie braucht nicht bei Verwendung von $re(r!)$ postuliert zu werden.) - Die gefundene symmetrische Wirksamkeit von r und f unter der akt-2- (und nreh-2-)Bedingung spricht gegen die Allgemeinheit des "asymmetrischen Gesetzes des Effekts" nach THORNDIKE.

Für allgemeine k=4-Bedingungen reicht die rein kognitive Informationshypothese (plus Strategieneklärung) nach BUCHWALD nicht aus. Wie auf S. 135f. dargestellt, gilt sie zunächst für die BUCHWALDsche Versuchsanordnung einschließlich der von ihm

benutzten Instruktion, die auf Treffermaximierung abzielt (s. Punkt 2). Ändert man die Instruktion in Richtung Symmetriebetonung ab (s. Punkt 3), so müßte ein ebenso rein kognitiver Standpunkt (bei entsprechend rational geänderter Strategie) Symmetrie voraussagen. Diese wurde nun aber nur unter beob-Bedingungen gefunden. Zur Interpretation der asymmetrischen Ergebnisse aus akt-4-Situationen muß man den rein kognitiven Standpunkt verlassen, der eine vollbewußte Informationsausnutzung beinhaltet, zugunsten einer Erklärung mit Hilfe des postulierten automatisierten Kontrollprozesses. Dieser wurde zwar im Rahmen eines kognitiv ausgerichteten Gedächtnismodells eingeführt (s. 6.3.3), unterscheidet sich aber deutlich von der "Kognitivität" der Informationshypothese, wie sie üblicherweise dargestellt wird.

Die obigen Ergebnisse stellen einen neuen Aspekt in der Behandlung der relativen Effizienz von r und f in $k=4$ -($k>2$ -) Situationen heraus. Nach SCHWARZE (1973, S. 49) dagegen war "die Verwendbarkeit von Mehrfachwahlsituationen zur Überprüfung der relativen Effizienz von Lob und Tadel grundsätzlich in Frage zu stellen."

Es soll an dieser Stelle eine kleine Spekulation über den Vorgang der "Symmetrisierung" beim Übergang von der akt-4-zur beob-4-Bedingung eingeschoben werden (s. S. 211). Die Argumentation war gewesen, daß aufgrund des automatisierten Kontrollprozesses unter akt-4-Lernbedingungen die Aufmerksamkeitszuwendung den f -Items gegenüber relativ gering sei zugunsten der r -Items. Wird nun beim Übergang zur symmetrischen beob-4-Bedingung die gleichmäßige Aufmerksamkeitszuwendung das hohe Niveau bezüglich der akt- r -Items haben, oder wird sie auf das niedrige Niveau bezüglich der akt- f -Items abgesenkt werden, oder wird sich ein mittleres Niveau einspielen? Die letzte Möglichkeit ist die wahrscheinlichste: Sie drückt eine Art von "Erhaltungssatz für die Summe der Aufmerksamkeitszuwendungen" aus.

Dieser Gedanke soll numerisch verfolgt werden, und zwar an Hand der Daten in der oberen Hälfte von Tab. 7.1.3-1 auf S.

194; als Maße für "Aufmerksamkeitszuwendung gegenüber r- bzw. f-Items" werden also die relativen Häufigkeiten vorgeschlagen, mit denen r bzw. f w i r k s a m geworden sind. Zu vergleichen sind demnach die Werte $n'(R-R!r!)$ und $n'(\bar{R}-R!f!)$ aus der akt-4-Bedingung mit denen aus der beob-4-Bedingung.

Zuvor muß aber noch ein Korrekturfaktor bestimmt werden, der die offensichtlich allgemein niedrigere Aufmerksamkeitszuwendung unter beob-Bedingungen in Rechnung setzt. Dieser Faktor wird zwischen den akt-2- und beob-2-Gruppen berechnet, da hier kein Wechsel der Symmetrieverhältnisse stattfindet:

$$(17,8\%+16,1\%)/(25,6\%+27,1\%)=0,64.$$

Mit diesem Wert ist die Gesamtaufmerksamkeitszuwendung ($=LE'$; s. S. 193) der akt-4-Gruppe zunächst zu korrigieren:

$$0,64 \cdot (28,4\%+11,9\%)=18,4\%+7,6\%=26\%.$$

Die asymmetrische Zusammensetzung von 26% aus 18,4% und 7,6% sollte in der beob-4-Gruppe in eine symmetrische übergehen, nämlich in 13% + 13%. Die tatsächlichen Werte sind 13,3% und 15,3%. In Anbetracht der Schwankungsmöglichkeiten zwischen den vier beteiligten Gruppen ist der obige "Erhaltungssatz" befriedigend bestätigt.

Die Berechnung der ϕ -Koeffizienten erbrachte durchschnittlich zufriedenstellende Resultate. Sie diente im wesentlichen zur zusätzlichen Illustration der Verhältnisse, wie sie nach dem modifizierten BM (s. 6.3.1) auftreten können, z.B. zur Demonstration der Ungültigkeit von U_3 (d.h. $\phi=0$), wenn Rehearsal erlaubt ist. Für die Behandlung der Latenzzeiten in vorliegender Arbeit können die ϕ -Koeffizienten keine direkte Hilfe leisten. Bezüglich einer eventuellen Weiteranalyse, z.B. von Einzelprotokollen, macht das Ergebnis $\phi > 0$ die Einschränkung, daß die Erinnerungszeiten an R und O längs eines (c,1-u)-Weges nicht als voneinander unabhängig anzunehmen sind; es entfiele daher die Möglichkeit, die Verteilung der zusammengesetzten Zeiten mit Hilfe von Momenterzeugenden Funktionen zu behandeln (z.B. ATKINSON et al., 1965, S. 168f.).

Die Ergebnisse bezüglich des Vergleichs der Variationsbreiten von $re(b)$ und $re(r!)$ zugunsten des letzteren Effizienzmaßes

sprechen für sich (s. S. 195f.).

Die Bedeutung hinsichtlich von $re(b)$, aber auch die Einschränkungen der Berechnung der Beharrungsparameter g' und g'' wurden im Ergebnisteil auf S. 196ff. schon ausführlich behandelt. Es soll nicht mehr weiter darauf eingegangen werden, da im Augenblick der Schwerpunkt der Diskussion auf dem relativen Effizienzmaß $re(r')$ liegt, bei dem eben der g -Mechanismus keine Rolle mehr spielt. Nichtsdestoweniger sollten bei Untersuchungen, die die konventionellen Wiederholungsdaten gebrauchen, die Überlegungen zur k -Abhängigkeit von g bzw. $re(b)$ im Auge behalten werden.

Die Interpretation der a-posteriori-Wahrscheinlichkeiten in den asymmetrischen Gruppen (akt-4, nreh-4; s. Abb. 7.1.3-2 auf S. 198) bereitet Schwierigkeiten: Es ist plausibel, daß eine geringere Aufmerksamkeitszuwendung gegenüber f -Items zu einer geringeren $O=f$ -recall-Leistung führt, wie es auch gefunden wird (s. Bild (K) und (M) auf S. 189), die Anwendung der Kategorie "sicher" sollte deshalb aber nicht inadäquater als bei r -Items vorgenommen werden.

Eine mögliche Interpretation wäre: Unter "aktiven" $k=4$ -Bedingungen konzentrieren sich die V_{pn} gemäß dem angenommenen automatisierten Kontrollprozeß im Lerndurchgang besonders auf r -Items. Wenn sie sich darauf im Prüfdurchgang an r , f und b erinnern sollen, kann ihnen nachträglich wieder einfallen, daß im Lerndurchgang f doch einigermaßen häufig aufgetreten war, sie sich aber nicht entsprechend auf die f -Items konzentriert hatten. Dieses kann dazu führen, daß die V_{pn} im Falle von tatsächlichen Nichterinnerungen auf Verdacht hin " f !" angeben, um beim V_l einen "guten Eindruck zu machen". Diese Strategie ist nämlich nicht ganz uneffektiv, da unter den Nichterinnerungen sich hauptsächlich f -(und b -)Items befinden werden; die r -Items werden ja gut erinnert.

Die Beiträge für die fälschlichen " f !"-Erinnerungen werden sich daher zum großen Teil aus tatsächlichen b -Rückmeldungen rekrutieren, weniger aus r -Rückmeldungen. Diese letztere Vermutung wird bei genauerer Inspektion der Daten zur Berechnung

der a-posteriori-Wahrscheinlichkeiten (hier nicht angegeben) leicht unterstützt.

Bei der Darstellung der Ergebnisse bezüglich der Latenzzeitmessungen wurde die Diskussion schon teilweise vorweggenommen. Die im wesentlichen positive Beantwortung des Hypothesenkomplexes H1 bis H9 weist auf die Konsistenz der Überlegungen zu den Latenzzeiten mit den Annahmen des modifizierten BM hin (die endgültige Diskussion wird noch auf 7.3 und 8.1 verschoben). Insbesondere konnten auch die Zwischenschritte, die bei der Ableitung der Hypothesen gemacht wurden, teilweise demonstriert werden, z.B. die Isolierung der "sicheren, schnellen" (c,1-u)-Zeiten nach Anteil und Größe (s.S. 205ff.). Es wurde aber auch darauf hingewiesen, daß unter Umständen der Anteil "schneller" Zeiten zu gering sein könnte, um einen beobachtbaren Effekt zu erzielen.

Die Fragestellung (3), zu der die Hypothese H10 aufgestellt wurde, muß in gewissem Sinne als negativ beantwortet gelten. Zwar waren alle Voraussetzungen für ein positives Resultat gegeben, d.h. die Fragestellungen (2) und (5a) waren hinreichend bestätigt, und auch die speziellen Voraussetzungen für einen Effekt entsprechend H10 konnten isoliert als erfüllt demonstriert werden (s.S. 209), der tatsächlich beobachtete Effekt war aber nur schwach unterstützend. Es wurde gefolgert, daß sich zwar die Einzelüberlegungen zu H10 bestätigt hatten, daß aber die Zufallsschwankungen innerhalb und zwischen den beteiligten Gesamtverteilungen zu groß wären, um einen deutlichen Effekt nach H10 sichtbar werden zu lassen. - Dieses trifft wohl auch für die anderen Latenzzeitmuster zu, die sich im Zusammenhang mit dem Hypothesenschema auf S. 165 vergleichen lassen (s.S. 209).

Die Zufallsschwankungen werden dadurch so störend, daß sie sowohl über die Schwankungen der (c,1-u)- und GK-Mediane auf der Zeitachse als auch über die Schwankungen im prozentualen Mischungsverhältnis der "schnellen" und "langsamen" Zeitannteile wirksam werden (s. dazu S. 51!).

Die Fragestellung (3) kann nun nicht dadurch günstiger beantwortet werden, daß man einen präziseren Versuchsplan und eine genauere Auswertung anstrebt. Eine positive Beantwortung von (3) bestünde ja darin, daß der Effekt nach H10, nachdem er mit Hilfe der Verbalisierungsmethode begründet wurde, auch in nver-Situationen zu Aussagen über die relative Effizienz von r und f benutzt werden könnte, die sich aus den Wiederholungsdaten nicht oder nur unzuverlässig machen ließen. Dazu hätte aber der Unterschied zwischen den Differenzen $T_{ff} - T_{rr}$ aus den akt-4- und beob-4-Gruppen sehr viel ausgeprägter sein müssen (s. S. 172).

Zunächst wird also keine Möglichkeit zur positiven Beantwortung der Fragestellung (3) gesehen. -(Vielleicht ließe sich aus den sechs Medianen $T_{rr} \dots T_{bb}$ ein bezüglich (3) psychologisch sinnvoll gewichtetes Mittel berechnen, in dem sich die Schwankungen der Einzelmediane ausgleichen; oder es könnte das gesamte Latenzzeitmuster zur Differenzierung dienen und nicht nur die Differenz $T_{ff} - T_{rr}$ allein.)

Der postulierte "Zeitlupeneffekt" wurde zufriedenstellend bestätigt (in den weiteren Diskussionen in 7.3 und 8.1 kann daher von einer gewissen "Homogenität" der Latenzzeitmuster unter Verbalisierungs- und Nichtverbalisierungsbedingungen ausgegangen werden).

Für den Vergleich der nver- mit den akt-Bedingungen sollte aber eine sich andeutende kleine Diskrepanz im Auge behalten werden, die aus den vorliegenden Daten nicht näher zu analysieren ist (gegebenenfalls mit erheblich mehr Auswertungsaufwand, z.B. Analyse von Einzelprotokollen mit genauer Kontrolle der Strategien): Hinsichtlich der Wiederholungsprofile (s. Bild (G) auf S. 189) scheinen die beiden Gruppen unter der k=2-Bedingung sich mehr zu unterscheiden, hinsichtlich der Latenzzeitmuster mehr unter der k=4-Bedingung (s. S. 204f.).

Es sollen nun die unter 7.2 zu referierenden Ergänzungsuntersuchungen begründet werden.

Stellt man alle Ergebnisse zu den nreh-Gruppen zusammen und vergewärtigt man sich das experimentelle Vorgehen (s. S.

181), so liegt die Vermutung sehr nahe, daß die Verhinderung von Rehearsal mißlungen ist. Teilweise kann man daher die nreh-Gruppen als Replikationen der akt-Gruppen auffassen, teilweise muß man aber damit rechnen, daß durch die besondere Zeitabfolge im Lerndurchgang, speziell das lange Intervall zwischen R und O, möglicherweise Effekte mit ins Spiel kommen, die theoretisch irrelevant sind, jedenfalls im Rahmen dieser Arbeit. - Es soll daher noch ein Versuch unternommen werden, die Verhinderung von Rehearsal möglichst sauber durchzuführen (s. 7.2.1: Experiment IV).

Der Ausgangspunkt für das Experiment V (s. 7.2.2) war die Beantwortung der Frage F5 durch $k=2$ -Vpn (s. S. 210). Das Umschalten im Lerndurchgang von der Sequenz S-R-f auf S- $\bar{R}$ und das Abspeichern von S- $\bar{R}$ ist vollkommen im Einklang mit der rein kognitiven Informationshypothese bezüglich $k=2$ -Aufgaben. Falls aber das Umschalten auf S- $\bar{R}$ vollständig ist, müßte das einen Einfluß auf die Hypothese $H3: T_{rr} < T_{ff}$ (s. S. 71f.) haben. Im Prüfdurchgang kann die Vp auf S sofort mit $\bar{R}$ reagieren, ohne noch umschalten zu müssen (das ist ja schon im Lerndurchgang geschehen). Für diese "S- $\bar{R}$ "-Strategie wäre also $T_{rr} = T_{ff}$ vorauszusagen.

Betrachtet man in Abb. 7.1.3-3 auf S. 201 die Mediane T_{rr} und T_{ff} der akt-2- und nver-2-Bedingungen, so müßte bei der nver-2-Gruppe, in der diese besondere Strategie von 67% der Vpn angegeben war, eine überproportionale Annäherung von T_{ff} an T_{rr} zu beobachten sein, relativ zu den Verhältnissen in der akt-2-Gruppe. Ein solcher Effekt tritt nun allerdings nicht auf. Es soll im Experiment V abgeklärt werden, ob die Interviewdaten zu unzuverlässig sind oder ob das Umschalten von S-R-f auf S- $\bar{R}$ nicht so vollständig ist, daß daraus $T_{rr} = T_{ff}$ folgen würde.

Noch einige abschließende Bemerkungen zu den Interviewfragen (s. S. 210), soweit sie nicht schon anderweitig diskutiert wurden.

Zu F3: Erwartungsgemäß liegt der Prozentwert der beob-4-Vpn

näher bei dem der $k=2$ -Vpn als bei dem der "aktiven" $k=4$ -Vpn. Ob man den Unterschied zwischen 17% und 27% interpretieren kann, ist angesichts der auffallenden, aus objektiveren Daten gewonnenen Symmetrie in der beob-4-Gruppe fraglich.

Zu F4: Die Beantwortung dieser Frage ist ein Hinweis auf die schon öfter angesprochene Problematik des Einführens von "Verbalisieren". Demnach wurden Vpn aus Verbalisierungsgruppen stärker als nver-Vpn veranlaßt, auch b-Items zu lernen, obwohl ausdrücklich betont worden war, daß b-Wörter keinen Beitrag zum Lernerfolg liefern würden. Einige Vpn gaben explizit an, sie hätten nur wegen der Verbalisierungsanweisung sich auch auf b-Items konzentriert. - Diese Beantwortung von F4 ist allerdings nicht sehr schwerwiegend, da einerseits für die Berechnung von $re(r!)$ unter Verbalisierungsbedingungen die Einführung von b unnötig ist und andererseits in "konventionellen" nver-Experimenten diese Problematik überhaupt nicht auftritt. Nur wenn man am Vergleich von $re(b)$ unter ver- und nver-Bedingungen interessiert sein sollte, ergeben sich in dieser Hinsicht Schwierigkeiten.

7.2 Ergänzungsuntersuchungen

Die beiden folgenden Experimente wurden von einer studentischen Hilfskraft als V1 durchgeführt¹⁾. Durch intensive Einweisung des V1 durch den Vf. wurde versucht, die Versuchsbedingungen mit denen im Hauptexperiment (III) bis auf die beabsichtigten experimentellen Manipulationen möglichst vergleichbar zu machen. Ungeändert blieben Versuchsraum und -apparatur. Die Hauptinstruktion aus Experiment III fand weiterhin Verwendung. Es wurde nur "aktiv" gelernt. Vpn waren Studierende der Psychologie des 1. bis 4. Semesters, die zur Erlangung ihrer Pflicht-Vp-Stunden an den Experimenten teilnahmen. Keine der Vpn hatte schon an früheren Experimenten zur relativen Effizienz von r und f teilgenommen.

1) Ich danke Herrn K. Dittmann für die Ausdauer und Sorgfalt bei der Durchführung der Experimente IV und V.

7.2.1 Experiment IV: Striktes Nonrehearsal (k=4)

Die Versuchsbedingung "striktes Nonrehearsal unter k=4" wird im folgenden mit "nnreh-4" abgekürzt werden; diese Kurzbezeichnung soll auf das, relativ zur nreh-Bedingung, verstärkte Bemühen, Rehearsal zu unterbinden, hinweisen.

Fragestellung und Hypothesen

Die allgemeine Fragestellung wurde auf S. 218f. begründet. Aus zeitökonomischen Gründen sollte nur k=4 untersucht werden. Die k=4-Bedingung ist theoretisch relevanter als die k=2-Bedingung, weil sich bei erschwertem Rehearsal ein etwaiger Trend in Richtung auf das BM in einer Abnahme von $re(r!)$ bemerkbar machen müßte. Die speziellen Hypothesen sind also dieselben, wie sie zur nreh-4-Bedingung aufgestellt wurden (s. S. 166-168): 1. Annäherung von $re(r!)$ an 50%; 2. Annäherung von ϕ_r und ϕ_f an $\phi=0$; 3. Allgemeine Absenkung des Lernniveaus.

Methode

Die (einzige) Versuchsgruppe bestand aus 12 Vpn (4m;8w); das Alter der Vpn lag zwischen 18 und 25 Jahren. Die Versuchsdurchführung war die gleiche wie unter akt-4- bzw. nreh-4-Bedingungen (s. S. 177f.). Nur der Zeitablauf im Lerndurchgang war geändert, genauer: gestrafft. Im Vergleich zu den auf S. 181 dargestellten Abfolgen war der Zeitablauf für ein Item in diesem Experiment folgender (eckige Klammern bedeuten wieder Dias):

nnreh: <WIND>..5; Druck ; <GOLD> usw.
5 notieren
richtig

Die simultanen Tätigkeiten des V1 während der Transportzeit des nächsten Dias (Druck: "5 notieren" und "richtig") erforderten hohe Konzentration, konnten aber nach einigem Training gut bewältigt werden (die Antwortzeiten im Lerndurchgang wurden nicht mehr im Protokoll festgehalten). Durch dieses Vorgehen wurde also wieder die Zeit zwischen O und S minimiert, aber gleichzeitig auch die lange Pause zwischen R und O, wie unter der nreh-Bedingung, vermieden.

Die Auswertung folgte der des Experiments III. Zusätzlich sollte geprüft werden, ob unter diesen straffen Zeitbedingungen überhaupt ein Lernerfolg, gemessen durch die Steilheit des Wiederholungsprofils bzw. die Differenz rr-ff, aufgetreten war. Bei nicht vorhandenem Lernerfolg könnte sich nämlich $re(r!) \approx 50\%$ ergeben, wenn sowohl bei r- als auch f-Items gleichmäßig geraten würde. Dieses Ergebnis sähe dann fälschlicherweise nach einer Bestätigung des BM aus.

Ergebnisse

Die Ergebnisse unter der nnreh-4-Bedingung sind hauptsächlich von Interesse im Vergleich mit den Resultaten des Experiments III. Sie sind deshalb dort dargestellt:

S. 189: Abb. 7.1.3-1, Bilder (F), (J) und (M); der Übungseffekt (Position I vs. II) war erwartungsgemäß, ist aber nicht gesondert aufgezeichnet.

S. 194: Tab. 7.1.3-1; Tab. 7.1.3-2 wurde schon auf S. 195f. diskutiert; in Tab. 7.1.3-3 bringen die $g'-(g'-)$ Werte der nnreh-4-Bedingung keine zusätzlichen Einsichten zur Diskussion auf S. 196f.

S. 198: Abb. 7.1.3-2.

S. 200-203: Latenzzeitergebnisse.

Das Profil der Wiederholungshäufigkeiten in Bild (F), S. 189, ist relativ flach gegenüber den akt-4- und nreh-4-Profilen. Die Prüfung des Unterschieds zwischen rr und ff auf Signifikanz wurde mit Hilfe des WILCOXON-Tests durchgeführt (s. SACHS, 1969, S. 312; $\hat{R}=15,5 > R(11;5\%)_{\text{eins.}}=13$, also Nichtsignifikanz).

Trotz der nichtsignifikanten Lernleistung soll in der Darstellung der Ergebnisse weiter fortgefahren werden. Vergleicht man nämlich die akt-4-, nreh-4- und nnreh-4-Profile in Bild (F) miteinander, sieht man, daß sie erwartungsgemäß mit vermindertem Rehearsal flacher werden, aber weiterhin monoton abfallen. Es scheint nicht unvernünftig, auch der nnreh-4-Gruppe eine, wenn auch sehr geringe Lernleistung zuzuschreiben, die allerdings aufgrund des geringen Stichprobenumfangs inferenzstatistisch nicht nachzuweisen ist. (Die Anwendung

des etwas schärferen t-Tests für abhängige Stichproben würde kaum die weiteren Schlußfolgerungen beeinflussen.) Die folgenden Ergebnisse sind aber auf jeden Fall mit Vorsicht zu interpretieren.

Die statistische Prüfung von $re(r!) = 65,9\%$ (s. Bild (J), S. 189) auf bedeutsame Abweichung von $re(r!) = 50\%$ erbrachte Nichtsignifikanz (WILCOXON-Test: $\hat{R} = 18,5 > R(9; 5\%)_{\text{eins.}} = 8$).

Diskussion

Dieses letzte Ergebnis scheint für das BM zu sprechen, das bei verhindertem Rehearsal $re(r!) = 50\%$ voraussagt (s. S. 167). Vergleicht man aber auf S. 189 in den Bildern (H) bis (J) das nreh-4-Profil einerseits mit den akt-4- und nreh-4-Profilen, andererseits mit den vier übrigen Profilen, die knapp um $re(r!) = 50\%$ schwanken, so kann man allerdings nur von einer schwachen Tendenz in Richtung auf die Vorhersage des BM sprechen.

Auch in Tab. 7.1.3-1 auf S. 194 findet man einen ähnlichen Trend: Die $\emptyset$ -Koeffizienten nähern sich andeutungsweise der BM-Voraussage $\emptyset = 0$.

Aus dieser Tabelle und Tab. 7.1.3-2 können Werte entnommen werden, mit denen ein Vergleich der akt-4-, nreh-4- und nnreh-4-Gruppen (geordnet also nach verstärktem Nonrehearsal) hinsichtlich des auf S. 193 definierten Lernerfolgs $LE' = n'(R-R!r!) + n'(R-R!f!)$ und der relativen Effizienz $re(r!)$ möglich ist. In dieser Reihenfolge ergibt sich für LE' : 40,3%, 33,7%, 14,3% und für $re(r!)$: 70,4%, 69,1%, 65,9%. Mit verstärkten Bemühungen, Rehearsal zu verhindern, nimmt also der Lernerfolg LE' rapide ab, während die relative Effizienz $re(r!)$ sich nur wenig auf den vom BM geforderten Wert $re(r!) = 50\%$ hinbewegt.

Die Folgerungen sind: Durch verstärktes Nichtrehearsal ist zwar eine gewisse Annäherung an die Vorhersagen des BM zu erzielen, extrapolierend kann man aber annehmen, daß die Lernleistung bei noch größeren Anstrengungen, Rehearsal zu verhindern, auf uninterpretierbares Zufallsniveau absinkt,

bevor sich eine deutliche Bestätigung des BM ergibt (s. dazu 6.2.1).

Am Rande sei auf Bild "nnreh-4" in Abb. 7.1.3-2 (S. 198) hingewiesen, in dem sich eine gewisse Ähnlichkeit mit den asymmetrischen Gruppen (akt-4, nreh-4) zeigt (s. dazu die Diskussion auf S. 216). Im Verein mit dem numerisch relativ hohen Wert von $re(r!) = 65,9\%$ wird man die nnreh-4-Bedingung eher unter "asymmetrisch" als unter "symmetrisch" einzuordnen haben.

Bei den Latenzzeiten werden nach Tab. 7.1.3-5 auf S. 203 die Hypothesen H1 bis H6 (außer H5) befriedigend unterstützt. Im allgemeinen ist aber festzustellen, daß in den nreh- bzw. nnreh-Gruppen die Latenzzeithypothesen öfter verletzt sind als in den Gruppen mit erlaubtem Rehearsal. Es sieht so aus, als ob durch die Nonrehearsal-Bedingung schon während des Lerndurchgangs Streß bei den Vpn erzeugt wird, der zu einer hastigeren und ungeordneteren Einspeicherung und anschließenden Beantwortung führt.

Auch von dieser Seite her scheint es nicht besonders aussichtsvoll und theoretisch ertragreich zu sein, allzuviel Mühe auf das strikte Verhindern von Rehearsal zu verwenden (jedenfalls im Rahmen des augenblicklich behandelten Themas).

7.2.2 Experiment V: Strategiekontrolle ($k=2$)

Fragestellung und Hypothesen

Die allgemeine Fragestellung wurde auf S. 219 begründet. Es geht in diesem Experiment ($k=2$) darum, welche Auswirkungen es hat, wenn Vpn im Lerndurchgang auf die Sequenz S-R-f hin entweder mit Rehearsal dieser Sequenz S-R-f reagieren oder erst einmal auf S- $\bar{R}$ umschalten und dann die Sequenz S- $\bar{R}$ versuchen, sich einzuprägen. Diese unterschiedlichen Strategien (abgekürzt mit "S-R-f" bzw. "S- $\bar{R}$ ") sollen nicht nachträglich in einem postexperimentellen Interview erfragt und analysiert werden, sie werden vielmehr in zwei unabhängigen Versuchsgruppen von vornherein per Instruktion vorgegeben werden.

Selbstverständlich wird keine Verbalisierungsanweisung gegeben, weil sie mit großer Wahrscheinlichkeit die "S-R"-Instruktion bezüglich des Lerndurchgangs zunichte machen würde.

Da es vor allem auf die Auswirkungen der unterschiedlichen Strategien auf das Latenzzeitmuster ankommt, soll auch im Prüfdurchgang nicht nachträglich Verbalisieren eingeführt werden. Dieses könnte die beiden Gruppen unterschiedlich irritieren¹⁾.

Als spezielle Hypothesen werden aufgestellt:

$H(S-R-f): T_{rr} < T_{ff}$ (s. H3 auf S. 71f.)

$H(S-R) : T_{rr} = T_{ff}$ (s. S. 219).

Von Interesse ist auch das eventuell unterschiedliche Lernniveau der beiden Gruppen. Nach Bild (G) auf S. 189 deutete sich eine, zwar nicht signifikante, Überlegenheit der nver-2-Gruppe gegenüber der akt-2-Gruppe an (s. S. 190). In der ersteren hatten 67% der Vpn angegeben, die S-R-Strategie angewandt zu haben, gegenüber nur 15% in der akt-2-Gruppe (s. F5, S. 210).

In Experiment V wird "aktiv" und mit erlaubtem Rehearsal gelernt.

Methode

Die Versuchsgruppen bestanden unter beiden Strategiebedingungen aus je 11 Vpn (5m;6w). Das Alter der Vpn lag zwischen 19 und 35 Jahren.

Der Versuch begann zunächst wie in Experiment III unter der nver-2-Bedingung (s. S. 177-180). In der Hauptinstruktion

1) In diesem Experiment wurde tatsächlich auch ein, für die Vpn unerwarteter, z w e i t e r Prüfdurchgang mit Verbalisierungsanweisung durchgeführt, da aus den beiden Strategien unterschiedliche Vorhersagen bezüglich der Erinnerungsdaten abzuleiten sind. In diesem zweiten Prüfdurchgang aber lagen die a-posteriori-Wahrscheinlichkeiten so knapp über Zufallsniveau, daß nicht versucht werden soll, die Ergebnisse zu interpretieren, bevor nicht weitere Erfahrungen mit nachträglich im zweiten Prüfdurchgang eingeführten Verbalisierungsanweisungen vorliegen.

($k=2$) aber waren die letzten drei Zeilen ("Hast Du noch Fragen... beantworten.") abgedeckt, und es wurde fortgefahren mit der für die jeweilige Bedingung (S-R-f bzw. S-R̄) gültigen Strategieinstruktion (s. Anhang, S-R-f: S. 268; S-R̄: S. 269). Nach der Beantwortung von Verständnisfragen wurden dann Lern- und Prüfdurchgang (also ohne Verbalisieren) mit der Einübungsliste a' und anschließend mit der Hauptlernliste A' durchgeführt¹⁾. - Zum Abschluß des Versuchs wurden die Vpn gefragt, ob sie glaubten, sich an die für sie gültige Strategieanweisung gehalten zu haben.

Ergebnisse

Die Profile der Wiederholungshäufigkeiten sind in Abb. 7.2.2-1 dargestellt.

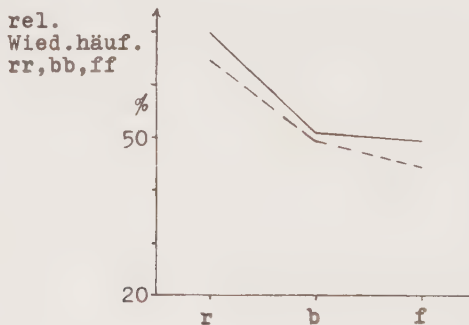


Abb. 7.2.2-1: Relative Wiederholungshäufigkeiten rr, bb und ff (in %); Strategie "S-R-f": ———; Strategie "S-R̄": -----; (Experiment V; $k=2$)

Der Lernerfolg ist geringfügig größer nach r für die S-R-f-Gruppe und nach f für die S-R̄-Gruppe. (Bei dem geringen Stichprobenumfang und den großen Streuungen lohnt es sich nicht, die sich andeutende Wechselwirkung zu testen.) Der gesamte Lernerfolg, ausgedrückt in absoluten Häufigkeiten der Wiederholungen nach r plus Nichtwiederholungen nach f, ist

1) Es erfolgte dann unmittelbar noch ein zweiter Prüfdurchgang zur Liste A' unter Verbalisierungsinstruktion; s. Fußnote auf S. 225.

für S-R-f: 159 und für S- $\bar{R}$: 158.

Auffallend sind die Werte bb , die für k=2-Bedingungen ungewöhnlich niedrig liegen.

Die Ergebnisse der Latenzzeitmessung sind in den Tabellen und der Abbildung auf S. 200-203 dargestellt. Die Hypothesen H1 bis H6 werden nach Tab. 7.1.3-5 befriedigend unterstützt (Ausnahme: H5 unter der S-R-f-Bedingung). Bezüglich der für dieses Experiment aufgestellten speziellen Hypothesen H(S-R-f) und H(S- $\bar{R}$) sind die Ergebnisse jedoch der Erwartung entgegengesetzt ausgefallen: S-R-f-Bedingung: $T_{rr}=2,15 \text{ sec} \approx T_{ff}=2,16 \text{ sec}$; S- $\bar{R}$ -Bedingung: $T_{rr}=2,04 \text{ sec} < T_{ff}=2,23 \text{ sec}$.

Diskussion

Die niedrigen bb -Werte sind vermutlich durch die zweimaligen Hinweise auf die Irrelevanz der Rückmeldung b in den Strategieinstruktionen zu erklären. Dieser Effekt war nicht beachtet, spielt aber für die weitere Diskussion keine Rolle.

Während man aus den Wiederholungsdaten der nver-2-Gruppe eine gewisse Überlegenheit der S- $\bar{R}$ - gegenüber der S-R-f-Strategie herauslesen konnte, findet man in diesem Experiment eine nahezu auffällige Übereinstimmung im Lernerfolg unter beiden Strategiebedingungen.

Obwohl alle Vpn im postexperimentellen Interview angaben, sich an die für sie gültige Strategie gehalten zu haben, bleiben mindestens zwei Interpretationen bezüglich des übereinstimmenden Lernerfolgs: Entweder haben die Vpn sich tatsächlich vollständig an die jeweilige Instruktion gehalten, dann spricht das Ergebnis für gleiche Effektivität der beiden Strategien, oder die Vpn waren nur subjektiv der Meinung, die ihnen vorgeschriebene Strategie befolgt zu haben, dann kann es sein, daß sie in Wirklichkeit eine aus "S-R-f" und "S- $\bar{R}$ " gemischte Strategie gebraucht haben; im letzteren Fall enthält die gefundene Übereinstimmung im Lernerfolg keine Aussage mehr.

(Natürlich kann die hundertprozentige Bejahung der Interviewfrage auch aufgrund "sozialer Erwünschtheit" zustande gekommen sein.)

Auf jeden Fall hat sich aus diesem Experiment keine eindeutige Überlegenheit der S- $\bar{R}$ - gegenüber der S-R-f-Strategie ergeben. Von diesem Befund aus betrachtet scheint der leicht höhere Lernerfolg der nver-2-Gruppe in Experiment III, verglichen mit der akt-2-Gruppe, nicht auf der bevorzugt angegebenen S- $\bar{R}$ -Strategie zu beruhen. Vielmehr kann man wohl als Erklärung einfach die geringere, also weniger verwirrende, Komplexität der Versuchssituation unter nver-Bedingungen annehmen.

Zu den unerwarteten Ergebnissen bezüglich der Latenzzeithypothesen H(S-R-f) und H(S- $\bar{R}$) soll eine noch spekulative Interpretation angeboten werden. Diese betrifft zunächst die Vpn unter der S- $\bar{R}$ -Anweisung: Wenn im Lerndurchgang die Sequenz S-R-f auftritt, bemüht sich die Vp, von S-R-f auf S- $\bar{R}$ umzuschalten und nur diese letztere Sequenz in den Langzeitspeicher zu transferieren. Wenn ihr dieses vollkommen gelingt, d.h. wenn die Vp bei Erinnerung an S- $\bar{R}$ nur noch per Zufall entscheiden kann, ob im Lerndurchgang S-R-f oder S- $\bar{R}$ -r aufgetreten war, dann würde tatsächlich als Vorhersage: $T_{rr} = T_{ff}$ folgen.

Vermutlich wird aber etwas anderes abgespeichert als nur die Sequenz S- $\bar{R}$. Es ist unwahrscheinlich, daß die Vp in der relativ kurzen Zeitspanne zwischen Auftreten des betrachteten Items im Lern- bzw. im Prüfdurchgang den bewußt vollzogenen Umschaltvorgang vergessen hat. Vielmehr ist anzunehmen, daß die Vp tatsächlich folgende Sequenz in den Langzeitspeicher transferiert hat: S-(R-f)- $\bar{R}$. Durch die Einklammerung von R-f soll symbolisch angedeutet werden, daß die Konzentration der Vp während des Transfers, nach ihrer subjektiven Meinung jedenfalls, nur auf $\bar{R}$ lag.

Wird nun im Prüfdurchgang wieder S dargeboten, so sieht sich die Vp mit den Erinnerungen an $\bar{R}$ und R-f konfrontiert. Dabei ist möglicherweise die Erinnerung an R-f durchaus schwächer als die an $\bar{R}$. Nichtsdestoweniger kann doch eine gewisse Zeitspanne vergehen, bis die Vp z.B. festgestellt hat, daß f zu R und nicht zu $\bar{R}$ gehört und daß sie eigentlich R-f gar nicht zu

beachten brauchte. Unter Umständen kann also der ganze Umschaltvorgang aus dem Lerndurchgang noch einmal ablaufen. Daraus würde wieder die Hypothese $H_3: T_{rr} < T_{ff}$ folgen.

Auch bei den Vpn unter der S-R-f-Anweisung gilt weiterhin H_3 . Es ist aber vorstellbar, daß diese Vpn bei Darbietung von S im Prüfdurchgang nur an R-f erinnert werden und in verhältnismäßig kürzerer Zeit als die S- $\bar{R}$ -Vpn mit $\bar{R}$ reagieren können.

Diese Überlegungen könnten also die oben erhaltenen Ergebnisse bezüglich T_{rr} und T_{ff} plausibel machen. Sie sind auch mit der zweiten Interpretationsmöglichkeit für den übereinstimmenden Lernerfolg unter beiden Strategiebedingungen in Beziehung zu setzen. - Es sei betont, daß die obige spekulative Betrachtung sich vollständig im Rahmen der rein kognitiven Informationshypothese für $k=2$ -Situationen bewegt.

Das Experiment V hat keine einigermaßen abschließende Lösung des speziellen Strategieproblems, "S-R-f" vs. "S- $\bar{R}$ ", in $k=2$ -Aufgaben gebracht. Es deutet sich aber die Möglichkeit an, mit Hilfe von Latenzzeitmessungen solche Probleme zu bearbeiten (s. dazu z.B. HERRMANN u. STAPF, 1973).

Für die vorliegende Arbeit ist das unter der S- $\bar{R}$ -Bedingung gefundene Ergebnis: $T_{rr} < T_{ff}$ positiv zu werten. Es bedeutet, daß die Hypothese H_3 in ihrer ordinalen Form für alle bisherigen Experimentalgruppen beibehalten werden kann, wenn auch die Begründung im einzelnen differenzierter geworden ist.

7.3 Gesamtauswertung zu den Latenzzeiten (Experimente I bis V)

Nachdem eine Reihe von Experimenten (I bis V) durchgeführt worden ist, in denen Latenzzeiten gemessen wurden, stellt sich das Problem der Auswertung in einer neuen Sicht. Im Zusammenhang mit Experiment I war in 5.2.2 auf S. 98ff. diskutiert worden, welche Schwierigkeiten eine saubere inferenzstatistische Auswertung der Daten wegen der unkontrollierbaren partiellen Abhängigkeit der Häufigkeitspolygone bringen würde.

Jetzt liegen die Latenzzeitmuster von 15 Versuchsgruppen vor, die voneinander unabhängig sind (s. Tab. 7.1.3-4, S. 200, Abb. 7.1.3-3, S. 201; auf die Abhängigkeitsverhältnisse in Experiment II wird noch einzugehen sein). Man kann diese Latenzzeitmuster als 15 unabhängige Schätzungen eines "wahren" Latenzzeitmusters betrachten (mit "Muster" ist immer die Rangreihenfolge der Mediane gemeint, nicht die absolute Lage auf der Zeitachse). A priori darf man nicht sagen, daß aufgrund der unterschiedlichen Versuchsbedingungen irgendeine Gruppe ein "richtigeres" Muster geliefert haben könnte als andere; das bezieht sich jedenfalls auf die Aspekte in den Mustern, soweit sie im Rahmen der Hypothesen H1 bis H9 erfaßt sind, die ja unabhängig von diesen Versuchsbedingungen aufgestellt wurden.

Anders ausgedrückt können die 15 unabhängigen Versuchsgruppen als Individuen einer Stichprobe aufgefaßt werden, von denen je 6 abhängige Meßwerte, nämlich die Mediane, vorliegen.

Bevor nun von einer Stichprobe mit 15 unabhängigen Individuen ausgegangen werden kann, muß noch etwas zum Verhältnis der beiden Bedingungen R-corr(A) und R-corr(B;speed) des Experiments II gesagt werden, die ja beide mit denselben Vpn durchgeführt wurden, insofern abhängig sind. Es soll für die weitere Auswertung festgelegt werden, daß diese beiden Bedingungen aufgrund des unterschiedlichen Übungsstadiums und vor allem der sie unterscheidenden "speed"-Instruktion als hinreichend unabhängig voneinander betrachtet werden können. Durch diese "Unabhängigkeitserklärung" soll einerseits ein Verlust an Information (durch Eliminierung einer der beiden Bedingungen) vermieden werden, dürften andererseits kaum ernste Verfälschungen hinsichtlich der weiteren Schlußfolgerungen entstehen.

Im Prinzip könnten nun mit Tab. 7.1.3-4 ein FRIEDMAN-Test durchgeführt und bei Signifikanz multiple Vergleiche nach WILCOXON u. WILCOX angeschlossen werden (s. SACHS, 1969, S. 527f.; S. 532f.; $n=15$ Individuen oder Blöcke, also die Versuchsgruppen, unter 6 Bedingungen). Eine Signifikanz im

FRIEDMAN-Test würde die Fragestellung (1) positiv beantworten, wäre dann aber für die weiteren Diskussionen nicht besonders hilfreich. Die multiplen Vergleiche sind aber a-posteriori-Tests und entsprechend unscharf: Sie berücksichtigen nicht die schon a priori aufgestellten Hypothesen H1 bis H9.

Es wird daher auf dieses formal korrekte Verfahren verzichtet zugunsten eines gemischt deskriptiv-inferenzstatistischen Vorgehens, das eine zweckmäßige Kombination von Anschaulichkeit (hinsichtlich der Einzelaussagen in H1 bis H9) und objektiver Schlußweise darstellen soll.

Im weiteren wird von Tab. 7.1.3-5 auf S. 203 ausgegangen. Hypothese H10 wurde schon auf S. 208f. endgültig diskutiert, H1 wird noch weiter unten behandelt werden.

Das deskriptiv-inferenzstatistische Verfahren besteht darin, die in der Tabelle enthaltenen Informationen bezüglich (Nicht-)Bestätigung der Hypothesen zu komprimieren, indem für jede der Hypothesen H2 bis H9 ein einseitiger Vorzeichentest gerechnet wird. Jede von ihnen wird dann nur noch durch die im Test erreichte Überschreitungs- (Irrtums-)Wahrscheinlichkeit $P(X \geq x/n = \dots; p=0,5)$ gekennzeichnet (s. KREYSZIG, 1965, S. 337f. u. 389); n ist für H2 bis H6 gleich 15, für H7 bis H9 gleich 5. Die "Treffer" x (=Anzahl der "+"-Zeichen) und die erhaltenen Wahrscheinlichkeiten sind in Tab. 7.1.3-5 in den beiden untersten Zeilen zu jeder Hypothese eingetragen.

Betrachtet man jede der Hypothesen für sich und vergleicht die erhaltenen Überschreitungswahrscheinlichkeiten mit $\alpha=5\%$, so sprechen die Ergebnisse (außer bei H8) sehr eindeutig für die Hypothesen. Mit diesem Schritt wird allerdings auch schon der Bereich des korrekten inferenzstatistischen Vorgehens überschritten, da sofort das Problem der Änderung des Signifikanzniveaus α bei Durchführung mehrerer Tests auftaucht und die Tests teils über den gemeinsamen Datensatz (z.B. H2 und H2' bezüglich T_{rr}), teils über die Struktur der Hypothesen voneinander abhängig sind. Z.B. ist aus den Eintragungen (+ bzw. -) unter H4' und H5 eindeutig vorherzusagen, daß

unter H2' nur "+"-Zeichen stehen werden; diese drei Hypothesen sind aber unabhängig voneinander begründet worden (s. S. 71f.) und sollen daher auch einzeln beurteilt werden können. Die Überschreitungswahrscheinlichkeiten in der letzten Zeile von Tab. 7.1.3-5 sind daher korrekterweise nur als deskriptive Zusammenfassung der jeweiligen Spalteninformation zu beurteilen; nichtsdestoweniger bleiben die niedrigen Werte beeindruckend. - Es ist darauf hinzuweisen, daß sich diese niedrigen Werte ergaben, obwohl ein relativ schwacheffizienter Test angewandt wurde, der viel an numerischer Information in den Daten verschenkt (zur Effizienz des Vorzeichen-tests s. z.B. SACHS, 1969, S. 317).

Aus eben diesem Hinweis ist auch zu begründen, warum die Hypothese H1: $T_{ff} = T_{rr}$ nicht mit dem Vorzeichentest geprüft wurde. Da es sich bei H1 um eine "zu bestätigende Nullhypothese" handelt, würde bei der Wahl eines relativ unscharfen Tests von vornherein zugunsten von H1 gearbeitet. Auch wegen der Bedeutung von H1 im Zusammenhang mit den Ableitungen der anderen Hypothesen soll daher der schärfere Test nach WILCOXON angewandt werden. Es ist dabei von den beiden Spalten (T_{ff}) und (T_{rr}) in Tab. 7.1.3-4 auf S. 200 auszugehen (s. SACHS, 1969, S. 312f.; $\hat{R}=32 > R(15;5\%)_{\text{zweis.}}=25$, d.h. Nichtsignifikanz). Der Ausgang des Tests spricht also nicht gegen die Hypothese H1: $T_{ff} = T_{rr}$.

Die Hypothese H8: $(T_{b5} - T_{bb})_{k=2} < (T_{b5} - T_{bb})_{k=4}$ ist die einzige der Hypothesen, bei der das Ergebnis der Erwartung nicht entsprach. Es ist zu vermuten, daß die Wirkung der Entscheidung, nach b die Antwort zu wiederholen bzw. nicht zu wiederholen, sehr viel ausgeprägter ist als der Effekt, der durch die k-Abhängigkeit beim Umschalten entsteht. - Damit wird zunächst die Behandlung des Aspekts an den Latenzzeitmustern, soweit er durch den a priori aufgestellten Hypothesenkomplex H1 bis H9 erfaßt wird, abgeschlossen.

Betrachtet man noch einmal Tab. 7.1.3-5, so fallen zwar Versuchsbedingungen auf, bei denen einige Hypothesen verletzt werden, kein Latenzzeitmuster von irgendeiner Gruppe scheint

aber ein direkter "Ausreißer" zu sein. Auch von den unterschiedlichen Versuchsbedingungen her besteht kein Anlaß, einige der Muster prinzipiell abzulehnen.

Für das weitere werden daher die erhaltenen Latenzzeitmuster als "homogen" im eben besprochenen Sinne angenommen.

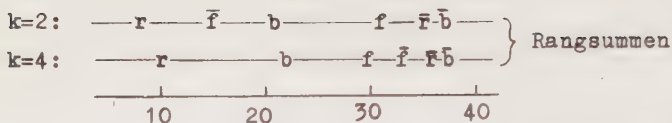
Es soll versucht werden, aus den 15 Mustern ein "mittleres" Muster zu berechnen, in dem die Beziehungen zwischen den Medianen als stabiler zu betrachten sind als in den individuellen (d.h. Gruppen-)Mustern. Diese Berechnung wird für die $k=2$ - und $k=4$ -Gruppen getrennt durchgeführt werden, da sonst der Effekt gemäß H7 und H8 verwischt würde.

Auszugehen ist von Tab. 7.1.3-4 auf S. 200. Zur Bestimmung des "mittleren" Musters ist es unvorteilhaft, für die sechs Spalten (T_{rr})..(T_{bb}) einfach die Mittelwerte zu berechnen. Dadurch blieben die Rangreihenfolgen von Gruppen mit kleiner Spannweite (z.B. $k=2$ -R-corr(I)) unter Umständen vollständig unberücksichtigt (s. zur Veranschaulichung Abb. 7.1.3-3).

Es soll daher so vorgegangen werden, daß alle Muster gleichberechtigt zum "mittleren" Muster beitragen, unabhängig von ihrer Spannweite und absoluten Lage auf der Zeitachse. Dabei wird der Formalismus des FRIEDMAN-Tests benutzt, insofern Lokationsunterschiede mit Hilfe von Rangsummen behandelt werden. In Tab. 7.1.3-4 werden in den einzelnen Zeilen (= Gruppen) die Ränge 1 bis 6 entsprechend der Größe der jeweiligen Mediane verteilt und, getrennt für $k=2$ - und $k=4$ -Gruppen, die Summen über die einzelnen Spalten gebildet. Nach diesem Verfahren ergaben sich folgende Rangsummen:

	T_{rr}	T_{ff}	T_{bb}	T_{ff}	T_{rr}	T_{bb}	
$k=2$	8	15	21	31	35	37	(aus 7 Gruppen)
$k=4$	10	33	22	30	36	37	(aus 8 Gruppen)

Im Sinne des FRIEDMAN-Tests beinhalten diese Rangsummen durchaus metrische Information und sollen daher auch graphisch, wie folgt, dargestellt werden; die Bezeichnungsweise ist dabei die in Abb. 7.1.3-3 auf S. 201:



Diese dargestellten Rangreihenfolgen sollen als "mittlere" Muster gelten.

(Die Muster für $k=2$ und $k=4$ kann man natürlich nicht absolut miteinander vergleichen.)

Die Beziehungen der Mediane innerhalb der "mittleren" Muster zueinander können als ziemlich stabil angesehen werden: Im $k=2$ -Muster sind die Daten von 123 Vpn, im $k=4$ -Muster die von 140 Vpn enthalten. Es ist auch auf die Übereinstimmung der Verhältnisse der Mediane untereinander sowohl im $k=2$ - als auch im $k=4$ -Muster hinzuweisen (eine Ausnahme macht der Median T_{ff} , bei dem sich die k -Abhängigkeit des Umschalteffekts erwartungsgemäß auswirkt).

Nach dem guten Abschneiden der Latenzzeithypothesen in den Einzelgruppen finden sich natürlich die Aussagen der Hypothesen H2 bis H8 auch in den "mittleren" Mustern wieder.

H9: $T_{k=2} < T_{k=4}$ kann nicht beobachtet werden, da durch die Rangbildung die Information über die Absolutzeiten verlorengeht; diese k -Abhängigkeit ist aber auch von relativ geringem Interesse. (Zur weiteren Diskussion, insbesondere bezüglich H1: $T_{ff} = T_{r\bar{f}}$, s. den folgenden Abschnitt 8.1.)

8. Schlußdiskussion

8.1 Zu den Latenzzeiten

Bevor die Bedeutung der Latenzzeitergebnisse für das Problem der relativen Effizienz von r und f näher diskutiert wird, soll noch eine nachträgliche Analyse angedeutet werden. Über die in Kap. 4 a priori abgeleiteten Latenzzeithypothesen hinaus kann versucht werden, aufgrund der auf S. 234 dargestellten "mittleren" Muster ad-hoc-Hypothesen aufzustellen, um Beziehungen zwischen den Medianen zu klären, die bisher noch nicht erfaßt wurden.

Wenn diese Beziehungen in den "mittleren" Mustern tatsächlich so stabil sind, wie auf S. 234 behauptet, dann liegt es z.B. nahe, den sowohl im k=2- als auch im k=4-Muster auffindbaren Unterschied zwischen T_{ff} und $T_{r\bar{r}}$ näher zu analysieren. Es ist also die augenscheinliche Beziehung $T_{ff} < T_{r\bar{r}}$ zu erklären. Diese Absicht steht nicht im Widerspruch zum nichtsignifikanten Ausgang des WILCOXON-Tests bezüglich H_1 , denn dadurch ist ja die strikte Beziehung $T_{ff} = T_{r\bar{r}}$ nicht bewiesen.

Da sich zunächst keine Erklärung anbot, warum T_{ff} kleiner als $T_{r\bar{r}}$ sein sollte, wurden für die willkürlich ausgewählte akt-2-Gruppe die Einzelprotokolle ausgewertet, d.h. für jede Vp ihre Zeitmediane bestimmt, insbesondere also T_{ff} und $T_{r\bar{r}}$. Von den sechs Einzelprotokollen, in denen sich die Beziehung $T_{ff} < T_{r\bar{r}}$ besonders deutlich zeigte, wurde eine "erweiterte" Strichliste angefertigt, d.h. anstelle einfacher Abzählstriche hinter den Zeitintervallen wurden die zu diesen Zeiten gehörigen Erinnerungsdaten notiert. Z.B. enthielt diese Strichliste bei den ff-Zeiten für das Zeitintervall 5,3-5,5 sec die Eintragungen: (5!b?), (3?f?), (5?b?), (5!r!). Die beiden Aufzeichnungen für die ff- und $r\bar{r}$ -Zeiten wurden nun darauf hin inspiziert, welche der Erinnerungskombinationen (R..O..) für die Beziehung $T_{ff} < T_{r\bar{r}}$ (teil-)verantwortlich sein könnte.

Nach Aufstellung, Nachprüfung und Verwerfung mehrerer Vermutungen schälte sich folgende Hypothese heraus: Im Lerndurchgang sei die Sequenz BERG 7 f erfolgt; die Vp habe sich aufgrund einer besonderen "Eselsbrücke" 7 sehr gut gemerkt, f dagegen etwas vernachlässigt und antworte im Prüfdurchgang auf BERG z.B. mit (7—7!r!). Die zugehörige Zeit ist also eine ff-Zeit, die aber wegen der subjektiven, wenn auch unbegründeten, Sicherheit relativ schnell erfolgen kann (s. S. 65). Die ff-Verteilung enthält dann außer den langsamen GK-Zeiten auch noch einen Anteil relativ schneller "falscher" (R-R!r!)-Zeiten, bei denen insbesondere auch kein Umschalten nötig war. Falls dieser Anteil groß genug ist, führt diese Überlegung zur Vorhersage: $T_{ff} < T_{r\bar{r}}$.

Wenn man diese "falschen" (R-R!r!)-Zeiten aus der ff-Verteilung herauszieht, müßte der Restmedian mit dem Median $T_{r\bar{r}}$ zusammenfallen oder sich ihm zumindest nähern; dieses wurde für die gesamte akt-2-Gruppe durchgeführt. Es ergab sich: Der Anteil der "falschen" (R-R!r!)-Zeiten an der ff-Verteilung war 12,5% mit einem Median von 3,27 sec. Der Restmedian betrug 5,72 sec, hatte sich also von $T_{ff}=5,54$ sec an $T_{r\bar{r}}=5,93$ sec angenähert.

Dieses Verfahren erhebt nicht den Anspruch, die Beziehung $T_{ff} < T_{r\bar{r}}$ vollständig erklärt zu haben. Es dient hauptsächlich zur Illustration, wie aufgrund der vorhandenen Zeit- und Verbalisierungsdaten durch sorgfältige (und zeitaufwendige) Inspektion ad-hoc-Hypothesen behandelt werden können. Offensichtlich wird durch den kleinen Effekt $T_{ff} < T_{r\bar{r}}$ die Begründung zur Hypothese H1 auf S. 71 etwas in Frage gestellt. Vermutlich ist die zu $T_{r\bar{r}}$ gehörige Verteilung allein als geeignetste Schätzung der GK-Verteilung anzunehmen.

Ein weiterer Anlaß für ad-hoc-Hypothesen wäre z.B. die Untersuchung von Gruppen, deren Latenzzeitmuster besonders stark von den "mittleren" Mustern abweichen; bei der Berechnung der "mittleren" Muster wurde eine Homogenitätsannahme bezüglich der individuellen (Gruppen-)Muster gemacht, die durch kein anerkanntes statistisches Verfahren abgesichert ist (wie z.B.

Prüfung auf Varianzhomogenität vor Rechnung einer Varianzanalyse).

Trotz dieser offenkundigen Lücken in der Behandlung der Latenzzeiten können die zu Beginn erhobenen Fragestellungen (1) und (2) als positiv beantwortet gelten. Die Latenzzeitmediane hängen nicht nur deutlich von den Rückmeldungsmodalitäten ab, ihre Beziehungen untereinander sind auch konsistent mit der gedächtnispsychologischen Interpretation des Verstärkungsvorgangs entsprechend dem modifizierten BM (s. 6.3.1 und die Ableitung der Hypothesen H1 bis H9 in Kapitel 4).

Diese Konsistenz ist insofern weitreichend, als die Hypothesen H1 bis H9 nicht nur durch die Ergebnisse der Verbalisierungsgruppen gestützt werden, sondern auch durch die der Gruppen, in denen nicht verbalisiert werden sollte (also die Gruppen der Experimente I, II und V und die unter nver-Bedingungen aus Experiment III). Bei diesen letzteren Gruppen entfällt der Einwand, daß durch die Verbalisierungsanweisung die gedächtnispsychologische Interpretation gemäß dem modifizierten BM praktisch erzwungen wurde und die Ergebnisse der Latenzzeitmessungen daher auch keine unabhängige Unterstützung dieses Aspekts liefern könnten.

Daran anschließend kann man die Informationshaltigkeit der Latenzzeitdaten so beurteilen: In den Verbalisierungsgruppen ist sie geringer als die der Verbalisierungsdaten selbst, die in diesen Experimenten die vergleichsweise präziseste und umfassendste Information überhaupt liefern. In den konventionellen Nichtverbalisierungsgruppen dagegen liefern die Latenzzeitdaten teilweise mehr Information als die reinen Wiederholungshäufigkeiten rr, bb und ff, da sich in den Zeitmustern gewissermaßen die gedächtnispsychologische "Erklärung" dieser Häufigkeiten widerspiegelt; dabei ist diese "Erklärung" über die Verbalisierungsdaten in den anderen Gruppen erprobt worden.

Eine Aussage der Latenzzeiten bezüglich der relativen Effizienz von r und f über die Aussage der konventionellen Häufigkeitsdaten hinaus (gemäß Fragestellung (3)) konnte nicht

positiv abgesichert werden: Die Information, die in den unterschiedlichen Differenzen $T_{ff} - T_{rr}$ aus den akt-4- bzw. beob-4-Gruppen durchaus erwartungsgemäß enthalten war, entsprach nur schwach (s.o.) der Information aus den zugehörigen Verbalisierungsdaten (s. S. 217); d.h. sie war zu unpräzise, um ohne gleichzeitige Erhebung von Verbalisierungsdaten reliable Schlüsse auf die relative Effizienz von r und f in spezifischen Situationen zuzulassen.

Wie auf S. 218 angedeutet, müßte vermutlich das gesamte Latenzzeitmuster zur Differenzierung zwischen zwei unterschiedlichen Versuchssituationen hinsichtlich der relativen Effizienz von r und f herangezogen werden. Dazu wäre es jedoch nötig, in den "mittleren" Mustern auf S. 234 nicht nur die Rangbeziehungen zwischen den Medianen, sondern auch die numerischen Abstände zwischen ihnen so genau zu verstehen, wie es nur aufgrund eines mathematisch formulierten Latenzzeitmodells möglich ist. - Die Konstruktion eines solchen Modells konnte in dieser Arbeit nicht geleistet werden. Die Begründung liegt in der Schwierigkeit, kein gut etabliertes und praktikables mathematisches Verstärkungsmodell zur Verfügung zu haben (s. S. 119 u. 133f.), und in der großen Streuung innerhalb der Latenzzeitverteilungen (s. S. 209).

8.2 Zum "Verbalisieren"

Im großen und ganzen kann gesagt werden, daß die Einführung des "Verbalisierens" im Prüfdurchgang (s. 6.3.4) sich für den vorliegenden Problembereich, also Latenzzeitmessung und relative Effizienz von r und f, bewährt hat. Insbesondere hat das vorgeschlagene relative Effizienzmaß $re(r!)$ eine Reihe von Vorteilen, die teilweise von vornherein ersichtlich waren, sich teilweise aber erst bei der Behandlung der Ergebnisse beweisen mußten.

Die Vorteile von $re(r!)$, besonders im Vergleich zu $re(b)$, seien noch einmal kurz zusammengestellt: Abhängigkeit von den Zufallsschwankungen nur z w e i e r Größen; keine problem-

irrelevante Abhängigkeit von k , z.B. über $bb=f(k)$; Unabhängigkeit vom Verständnis der Instruktion bezüglich der Bedeutung von b durch die V_{pn} ; Unabhängigkeit vom g -Mechanismus, der für das Problem der relativen Effizienz von r und f an sich irrelevant ist und nur durch die Erhebung von Wiederholungsdaten als abhängiges Maß ins Spiel kommt; als "symmetrisch" vorhergesagte Bedingungen (z.B. akt-2, beob-4,..) ergeben sehr stabile Werte mit geringer Schwankung um den Erwartungswert $re(r!)=50\%$; Anschaulichkeit und unkompliziertere Beziehung zum BM als $re(b)$; wegen der besonderen Bildung von $re(r!)$ aus Daten der Form $(R-R!r!)$ und $(\bar{R}-R!f!)$ wird die Problematik des abhängigen Lernmaßes (Wiederholungs- vs. Erinnerungsdaten) aufgehoben.

Unmittelbar an diesen letzten Punkt ist auch eine Kritik an $re(r!)$ anzuschließen, jedenfalls an der Form, wie dieses Maß hier erhoben wurde. Auf S. 159 (Nachteil Nr. 2) war gesagt worden, daß von $re(r!)$ nur im Rahmen einer gedächtnispsychologischen Interpretation des Verstärkungsvorgangs gültige Aussagen erwartet werden können, daß also die Anwendung dieses Maßes nur sehr eingeschränkt sein kann. Die vorliegende Untersuchung ist in der Tat von vornherein innerhalb eines kognitiven Bezugsrahmens konzipiert worden. (Sie stellt insofern eine Ergänzung oder auch einen Kontrast zu der Arbeit von SCHWARZE, 1973, dar, in der gerade über den Informationsgehalt hinausgehende Komponenten der Verstärkung analysiert werden.)

Der Informationsstandpunkt wurde allerdings in vorliegender Arbeit durch die Art der Prüfdurchgänge und den Zeitpunkt der Einführung der Verbalisierungsinstruktion stärker betont, als es z.B. in anderen Untersuchungen nötig wäre, in denen auch $re(r!)$ berechnet werden soll. Damit ist gemeint: In den Verbalisierungsgruppen der Experimente III und IV wurde schon in der Einübungsphase die Verbalisierungsinstruktion gegeben, die V_{pn} wußten also schon zu Beginn der Hauptlernlisten, daß sie im Prüfdurchgang nach ihren Erinnerungen an R und O gefragt würden, und konnten sich dementsprechend in den Lern-

durchgängen einstellen.

Diese experimentelle Abfolge von Verbalisierungsanweisung und Prüfdurchgängen ist natürlich nicht die einzig mögliche. Es gibt vielmehr eine Reihe von weiteren Kombinationsmöglichkeiten, z.B.: konventioneller Lern- und Prüfdurchgang, dann zweiter Prüfdurchgang mit R- und O-recall-Anweisung (s. z.B. POSTMAN u. ADAMS, 1954, 1955) oder: konventioneller Lerndurchgang, dann Prüfdurchgang nur unter R- oder O-recall-Anweisung (s. z.B. Experiment I) usw.

Die Kombination, die vermutlich für die theoretische Behandlung der relativen Effizienz von r und f am günstigsten ist und in der die Bestimmung von $re(r!)$ zu weiterreichenden Aussagen führen kann als in vorliegender Arbeit, ist folgende: Allgemeine Instruktion, Lerndurchgang, Prüfdurchgang mit für die V_p unerwarteter Verbalisierungsanweisung. - Es kann dann $re(r!)$, auch ϕ_r , ϕ_f usw., berechnet werden, wobei alle oben erwähnten Vorteile von $re(r!)$ erhalten bleiben. Es entfällt aber der Einwand der "im Lerndurchgang schon provozierten gedächtnispsychologischen Einstellung der V_{pn} ". Vielmehr ergibt sich die Möglichkeit zu untersuchen, wie starke Motivationen, die sich durch Art der Hauptinstruktion und des Lerndurchgangs erzeugen lassen, sich auf die einzelnen Aspekte der Verbalisierungsdaten auswirken können.

Nichtsdestoweniger bleibt auch so das Maß $re(r!)$ ein "gedächtnispsychologisches"; selbst unter der obigen günstigen Kombination gestattet es keine so weitreichenden Folgerungen, wie sie z.B. für die Behandlung des Problems andauernder unerwünschter Nebenerscheinungen von Bestrafungen wünschenswert wären.

Warum wurde nun diese günstige Kombination nicht in der vorliegenden Arbeit angewandt?

Naturgemäß kann die Verbalisierungsanweisung nur einmal unerwartet für die V_{pn} eingeführt werden, d.h. es kann nur eine Hauptlernliste unter diesen Bedingungen gegeben werden. Das Hauptziel vorliegender Arbeit war aber, den Beitrag von Ia-

tenzzeitmessungen zum Problem der relativen Effizienz von r und f zu untersuchen. Dazu mußten in einer Versuchssitzung pro V_p möglichst viel Zeitdaten erhoben werden, um einigermaßen stabile Aussagen bezüglich der Zeitmediane zu erhalten. Das wurde eben dadurch versucht zu erreichen, daß zwei Hauptlernlisten unter möglichst gleichen Bedingungen gegeben wurden.

Aus diesem Hauptziel erklären sich auch die Komplexität der Versuchsdurchführung und die Schwierigkeiten bei der statistischen Auswertung der Latenzzeitdaten. Mangels anderweitig erstellter und zur Verfügung stehender Daten mußten in *e i - n e r* Untersuchung Wiederholungshäufigkeiten, Latenzzeiten und die den Zusammenhang zwischen diesen beiden Datenarten transparent machenden Erinnerungsdaten erhoben werden. Das statistische Problem entsteht dadurch, daß pro V_p und Liste sich drei Wiederholungshäufigkeiten (rr , bb und ff) ergeben, denen 36 Zeiten gegenüberstehen, die ungleichmäßig auf sechs Kategorien (bezüglich $T_{rr} \dots T_{bb}$) verteilt sind. Die Häufigkeiten sind mit geläufigen statistischen Verfahren behandelbar, die Zeiten in der vorliegenden Form nicht.

8.3 Zur revidierten Informationshypothese

Dieser Abschnitt faßt nur kurz schon früher dargestellte, aber besonders wichtige Überlegungen zusammen.

Mit "revidierter" Informationshypothese ist die "klassische" Informationshypothese BUCHWALDS (1962) gemeint, zuzüglich der Einschränkungen, die durch Instruktionswechsel (z.B. "Symmetriebetonung"), unterschiedliche Lernmodalität (akt vs. beob) und den hypothetisch angenommenen automatisierten Kontrollprozeß bedingt sind. Diese Einschränkungen wurden schon auf S. 213f. hinreichend diskutiert. Es folgen einige weitere Anmerkungen.

Während der damaligen Behandlung (S. 148) nur beiläufig erwähnt, soll ein Aspekt der Überlegungen, die zur Hypothese $H(\text{akt}-4)$ führten, noch einmal herausgestellt werden. Bei den

bisher behandelten Experimentalsituationen handelt es sich um "vanishing situations" (s. S. 15), in denen die Möglichkeit der isolierten Erfassung des Effekts von f mit einer prinzipiellen Realitätsferne erkaufte wird. Aber gerade die Konfrontation eines im Alltagsleben unter der "Entdeckungs-Methode" erworbenen automatisierten Kontrollprozesses mit dieser artifiziellen Lernsituation unter der "Nichtkorrektur-Methode" führt zu einer Vorhersage bezüglich der Wirksamkeit von r und f , die eine gewisse Realitätsnähe beanspruchen kann, obwohl sie in einer künstlichen (dafür aber auch praktikablen) Laborsituation untersucht wird.

Ebenfalls gesondert soll noch einmal auf die nicht ganz eindeutige Interpretation des symmetrischen Ausgangs gemäß Hypothese H(beob-4) hingewiesen werden (s. S. 212f.). Insofern ist es nützlich, in den Fragestellungen (4) und (5) die praktische Herstellung von (a-)symmetrischen Situationen von ihrer theoretischen Abklärung abgetrennt zu haben. - Die weitere Bearbeitung des obigen Problems wird in vorliegender Untersuchung allerdings nicht mehr in Angriff genommen.

8.4 Kritik

Die Hauptkritik an dem experimentellen Vorgehen in dieser Arbeit ist gegen die Verwendung eines kleinen Antwortenrepertoires zu richten, z.B. 3 und 5 oder 3,5,7 und 9. Der Gebrauch von wenigen und gut bekannten Antwortalternativen ist eine vielgeübte und auch begründete Praxis in Lernexperimenten: Die Vpn brauchen z.B. nicht mehr die Antworten zu erlernen, die Analyse kann sich also auf den assoziativen Prozeß zwischen S und R konzentrieren. -

Wie aber die auf S. 173f. erwähnten 18 Protokolle, die eliminiert werden mußten, zeigen, besteht die Gefahr, daß die Vpn die Mehrfachwahlsituation in eine Begriffsbildungsaufgabe umwandeln (es jedenfalls versuchen) oder andere globale Lernstrategien entwickeln. Eine solche globale Lernstrategie wäre

z.B., im Lerndurchgang allen Reizwörtern mit den Anfangsbuchstaben a bis l eine 3 zuzuordnen, allen übrigen eine 5, und sich nur die Wörter zu merken, auf die der V1 "richtig" sagte; die Vp hätte sich also innerhalb einer Liste nur 12 Wörter zu merken.

Dieses Vorgehen stellt vom Standpunkt der Vp aus eine sehr effektive Strategie im Sinne des Lernerfolgs dar; damit wird aber, unabsichtlich, der Versuchszweck unterlaufen. Denn nach der Absicht des Experimentators sollte die Versuchsanordnung eine Vielzahl voneinander u n a b h ä n g i g e r Trial-and-error-Situationen anbieten, nämlich die einzelnen Reizwörter, auf die mit einer Reaktion aus dem vorgegebenen Repertoire zu reagieren ist.

Zur Abhilfe dieses unerwünschten Zustandes ist es allerdings problematisch, die Vpn zu deutlich auf die Vermeidung solcher Strategien hinzuweisen; es kann unter Umständen zur verstärkten Anwendung kommen, die dann aber im postexperimentellen Interview nicht zugegeben wird. - Selbst bei bewußtem Bemühen der Vpn, jedes einzelne Item für sich zu lernen, ist zu erwarten, daß unwillkürlich doch gewisse Gruppenbildungen und Verflechtungen stattfinden. Mit solchen "lokaleren", nicht genauer erfaßbaren Strategien ist sicherlich in den ausgewerteten Protokollen dieser Arbeit zu rechnen.

Naheliegend und auch wirksam wäre es, wieder auf die Versuchsanordnung THORNDIKES (1932) zurückzugreifen, wie sie auf S. 12 angedeutet wurde: Hier gehört zu jedem Reizwort ein gesondertes Alternativenrepertoire.

Die Erweiterung auf Erhebung von Zeit- und Verbalisierungsdaten böte kein Problem.

9. Zusammenfassung

Es wurde untersucht, welchen Beitrag Latenzzeitmessungen zum Problem der relativen Wirksamkeit positiver und negativer Rückmeldungen ("richtig"(r) bzw. "falsch"(f)) in Humanexperimenten zum verbalen PAL leisten können. Dabei wurde sich auf Alternativdiskriminationen ($k=2$) und Mehrfachwahlaufgaben ($k=4$) beschränkt.

Speziellere Fragestellungen waren: Hängen Latenzzeiten (insbesondere die Zeitmediane) überhaupt von den Rückmeldungsmodalitäten ab? Wenn ja, hängen sie von ihnen in einer Weise ab, die aufgrund theoretischer Vorstellungen über den Verstärkungsvorgang vorhersagbar ist? Erlauben Latenzzeiten Aussagen über die relative Effizienz von r und f, die über die aus Wiederholungsdaten gewonnenen Aussagen hinausreichen?

Diese Fragestellungen wurden im experimentellen und theoretischen Rahmen behandelt, der einerseits durch das "revidierte Gesetz des Effekts" von THORNDIKE und andererseits durch die "neue Interpretation" dieses Gesetzes mit Hilfe eines probabilistischen kognitiven Lernmodells (BM) von BUCHWALD (1969) aufgespannt wird. Insbesondere sollte das BM als Ausgangsbasis für ein mathematisches Latenzzeitmodell dienen, mit dem obige Fragestellungen zu behandeln wären.

Da es zu dem Thema nur wenige einschlägige Arbeiten gibt, wurden in einer breit gestreuten Literaturdurchmusterung meist mathematisch-psycho logische Latenzzeituntersuchungen aus verwandten Gebieten darauf hin geprüft, inwieweit sie empirische oder theoretische Hinweise zur Behandlung des gestellten Problems liefern könnten.

Für paarweise Beziehungen der Zeitmediane unter den verschiedenen Rückmeldungsmodalitäten (einschließlich einer "neutralen" Rückmeldung) wurde ein Hypothesenkomplex (H1 bis H9) abgeleitet, der sehr eng mit den Grundgedanken des BM zusammenhängt, aber nicht von der strikten Gültigkeit des BM abhängig ist.

Eine Voruntersuchung ergab, daß das BM in seiner strengen mathematischen Formulierung nicht als Ausgangsbasis für ein auf

ihm aufzubauendes Latenzzeitmodell geeignet sei. Die empirischen Zeitmediane unterstützten dagegen im wesentlichen den obigen auf einem abgeschwächten BM basierenden Hypothesenkomplex. In dieser Voruntersuchung führten technische Gründe zur Einführung von "Beobachtungslernen" (observational learning). Es wurde ferner die kritische Rolle diskutiert, die die strenge Verhinderung von Rehearsal spielt.

Eine Ergänzungsuntersuchung deutete auf bemerkenswerte Schwierigkeiten der Vpn hin, bei einer $k=4$ -Aufgabe sich ebenso sehr auf f -Items konzentrieren zu können wie auf r -Items. - Dieses Ergebnis und die genauere Diskussion der Informationshypothese BUCHWALDS (1962), die in gewissem Widerspruch zum BM von 1969 steht, führten zur Postulierung eines "stark automatisierten Kontrollprozesses". Dieser Kontrollprozeß sagt in $k>2$ -Situationen ähnliche Resultate voraus wie das "revidierte Gesetz des Effekts" von THORNDIKE, wird aber im Rahmen des kognitiven Gedächtnismodells von ATKINSON u. SHIFFRIN dargestellt.

Nach Verwerfung des strengen BM als Grundlage für ein Latenzzeitmodell wurde als neue Ausgangsbasis ein (Haupt-)Versuchsplan mit vier Gruppen aufgestellt, die sich nach Alternanzzahl ($k=2$ vs. 4) und nach Lernmodalität ("aktives" vs. "beobachtendes" Lernen) unterschieden. Aufgrund der um den "automatisierten Kontrollprozeß" und eine gewisse Instruktionsabhängigkeit erweiterten Informationshypothese wurden für drei dieser Gruppen Symmetrie, für die vierte Asymmetrie in der Effizienz von r und f vorhergesagt. Die Hoffnung war, daß sich diese unterschiedlichen Vorhersagen, falls sie bestätigt werden sollten, in den Latenzzeitmustern widerspiegeln würden. Eine diesbezügliche Hypothese (H_{10}) wurde mit Hilfe des abgeschwächten (modifizierten) BM abgeleitet. - In obigem Versuchsplan hatten die $k=2$ -Gruppen hauptsächlich Kontrollfunktionen.

Im Hauptexperiment wurden die Vpn aufgefordert, im Prüfdurchgang zu "verbalisieren", d.h. zu versuchen, sich an ihre Reaktionen und die Rückmeldungen im Lerndurchgang zu erinnern.

Durch dieses Verfahren sollte der Zusammenhang zwischen Wiederholungshäufigkeiten und Zeitdaten durchsichtig gemacht werden. Es wurde ein Maß für die relative Effizienz von r und f vorgeschlagen, das sich in vieler Hinsicht als vorteilhafter erwies als ein Maß, das auf der Anwendung einer "Nichts"-Rückmeldung basierte. - Die Einführung der Verbalisierungsmethode wurde mit Hilfe von Gruppen kontrolliert, in denen die Vpn nicht verbalisieren mußten.

Die Ergebnisse des Hauptexperiments unterstützten den Hypothesenkomplex (H1 bis H9) bezüglich der Zeitmediane, waren also konsistent mit einer gedächtnispsychologischen Interpretation des Verstärkungsvorgangs. Auch die Vorhersagen hinsichtlich Symmetrie und Asymmetrie im Hauptversuchsplan wurden deutlich bestätigt. Der Effekt gemäß Hypothese H10 war dagegen zu schwach ausgeprägt, um als relevant betrachtet werden zu können; für die schwache Ausprägung des Effekts sind die starken Schwankungen von Antwortzeiten verantwortlich zu machen, die bei Suchprozessen aus dem Langzeitspeicher entstehen.

In zwei Ergänzungsexperimenten wurde der Einfluß strikter Verhinderung von Rehearsal bzw. unterschiedlicher Strategien behandelt.

Die Verwendung eines kleinen, für alle Reizwörter einer Lernliste gemeinsamen Antwortenrepertoires wurde einer Kritik unterzogen.

10. Literaturverzeichnis

- ADAMS, J.A.: Human memory. New York: McGraw-Hill, 1967
- ADAMS, J.A., BRAY, N.W.: A closed-loop theory of paired-associate verbal learning. Psychol.Rev., 1970, 77, 385-405
- ALLEN, G.A., ESTES, W.K.: Acquisition of correct choices and value judgments in binary choice learning with differential rewards. Psychon. Sci., 1972, 27, 68-72
- AMSEL, A.: The role of frustrative nonreward in noncontinuous reward situations. Psychol.Bull., 1958, 55, 102-119
- Frustrative nonreward in partial reinforcement and discrimination learning. Psychol.Rev., 1962, 69, 306-328
- ARCHER, E.J.: A re-evaluation of the meaningfulness of all possible CVC trigrams. Psychol.Monogr., 1960, 74 (10, Whole No. 497)
- ATKINSON, R.C., BOWER, G.H., CROTHERS, E.J.: An introduction to mathematical learning theory. New York: Wiley, 1965
- ATKINSON, R.C., CROTHERS, E.J.: A comparison of paired-associate learning models having different acquisition and retention axioms. J.math.Psychol., 1964, 1, 285-315
- ATKINSON, R.C., ESTES, W.K.: Stimulus sampling theory. In: LUCE, R.D., BUSH, R.R., GALANTER, E. (Eds.): Handbook of mathematical psychology. Vol. II. New York: Wiley, 1963, 121-268
- ATKINSON, R.C., SHIFFRIN, R.M.: Human memory: a proposed system and its control processes. In: SPENCE, K.W., SPENCE, J. T. (Eds.): The psychology of learning and motivation. Vol. II. New York: Academic Press, 1968, 89-195
- ATKINSON, R.C., WICKENS, T.D.: Human memory and the concept of reinforcement. In: GLASER, R., 1971, 66-120
- AUDLEY, R.J.: A stochastic model for individual choice behavior. Psychol.Rev., 1960, 67, 1-15
- AUDLEY, R.J., MERCER, A.: The relation between decision time and relative response frequency in a blue-green discrimination. Brit.J. math.statist.Psychol., 1968, 21, 183-192

- AUDLEY, R.J., PIKE, A.R.: Some alternative stochastic models of choice. *Brit.J.math.statist.Psychol.*, 1965, 18, 207-225
- BANDURA, A.: Social learning through imitation. In: JONES, M.R. (Ed.): *Nebraska symposium on motivation: 1962*. Lincoln: Univ. of Nebraska Press, 1962, 211-269
- BANDURA, A., WALTERS, R.H.: *Social learning and personality development*. New York: Holt, Rinehart and Winston, 1963
- BARON, R.M.: Social reinforcement effects as a function of social reinforcement history. *Psychol.Rev.*, 1966, 73, 527-539
- BAUER, R., TEGEL, CH.: Einfluß der Rückmeldung beim Paarassoziationslernen auf Wiederholungshäufigkeiten und Latenzzeiten. Teil II. Unveröffentl. Semesterarbeit, Fachbereich Psychologie der Univ. Marburg, 1973
- BERGHÄUSER, K.-H., SCHRÖDER, G.: Unabhängige Erinnerungsprozesse für outcome und response beim Paarassoziationslernen. Unveröffentl. Semesterarbeit, Fachbereich Psychologie der Univ. Marburg, 1972
- BERGIUS, R.: *Psychologie des Lernens*. Stuttgart: Kohlhammer, 1971
- BERNBACH, H.A.: A forgetting model for paired-associate learning. *J.math.Psychol.*, 1965, 2, 128-144
- A multiple-copy model for postperceptual memory. In: NORMAN, D.A., 1970, 103-116
- BLÖSCHL, L.: *Belohnung und Bestrafung im Lernexperiment*. Weinheim: Beltz, 1969
- BOWER, G.H.: Choice-point behavior. In: BUSH, R.R., ESTES, W.K. (Eds.): *Studies in mathematical learning theory*. Stanford: Stanford Univer.Press, 1959, 109-124
- Application of a model to paired-associate learning. *Psychometrika*, 1961, 26, 255-280
 - An association model for response and training variables in paired-associate learning. *Psychol.Rev.*, 1962, 69, 34-53
 - A multicomponent theory of the memory trace. In: SPENCE, K.W., SPENCE, J.T. (Eds.): *The psychology of learning and motivation*. Vol. I. New York: Academic Press, 1967, 229-325

- BOWER, G.H., THEIOS, J.: A learning model for discrete performance levels. In: ATKINSON, R.C. (Ed.): Studies in mathematical psychology. Stanford: Stanford Univer.Press, 1964, 1-31
- BOWER, G.H., TRABASSO, T.R.: Concept identification. In: ATKINSON, R.C. (Ed.): Studies in mathematical psychology. Stanford: Stanford Univer.Press, 1964, 32-94
- BOX, G.E.P.: Non-normality and tests on variance. Biometrika, 1953, 40, 318-335
- BUCHWALD, A.M.: Extinction after acquisition under different verbal reinforcement combinations. J.exp.Psychol., 1959 a, 57, 43-48
- Experimental alterations in the effectiveness of verbal reinforcement combinations. J.exp.Psychol., 1959 b, 57, 351-361
 - Variations in the apparent effects of "right" and "wrong" on subsequent behavior. J.verb.Learn.verb.Behav., 1962, 1, 71-78
 - A family of mathematical models for the effectiveness of "right" and "wrong" in altering behavior. Tech.Rep. No. 2, NIMH M 5283 and MSP-18223. Indiana Univ., 1963
 - The effect of some variations in procedure on response repetition following verbal outcomes. J.verb.Learn.verb.Behav., 1966, 5, 77-85
 - Effects of immediate vs. delayed outcomes in associative learning. J.verb.Learn.verb.Behav., 1967, 6, 317-320
 - Effects of "right" and "wrong" on subsequent behavior: A new interpretation. Psychol.Rev., 1969, 76, 132-143
- BUSH, R.R.: Estimation and evaluation. In: LUCE, R.D., BUSH, R.R., GALANTER, E. (Eds.): Handbook of mathematical psychology. Vol. I. New York: Wiley, 1963, 429-469
- BUSH, R.R., MOSTELLER, F.: Stochastic models for learning. New York: Wiley, 1955
- BUSS, A.H., BRADEN, W., ORGEL, A., BUSS, E.H.: Acquisition and extinction with different verbal reinforcement combinations. J.exp.Psychol., 1956, 52, 288-295

- BUSS, A.H., WEINER, M., BUSS, E.H.: Stimulus generalization as a function of verbal reinforcement combinations. *J. exp.Psychol.*, 1954, 48, 433-438
- CAIRNS, R.B.: Informational properties of verbal and nonverbal events. *J.Pers.soc.Psychol.*, 1967, 5, 353-357
- CHRISTIE, L.S., LUCE, R.D.: Decision structure and time relations in simple choice behavior. *Bull.math.Biophysics*, 1956, 18, 89-112
- CLAUSS, G., EBNER, H.: Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen. Frankfurt a.M.: Deutsch, 1970
- COOMBS, C.H.: A theory of data. New York: Wiley, 1964
- COOMBS, C.H., DAWES, R.M., TVERSKY, A.: Mathematical psychology. Englewood Cliffs, N.J., Prentice-Hall, 1970
- DAND, A.: "Reward" and "punishment" in learning. *Brit.J.Psychol.*, 1946, 36, 83-87
- DENNEY, N.W.: Analysis of processing time for conjunctive and disjunctive problem solving. *Psychon.Sci.*, 1970, 20, 197-199
- EDWARDS, W.: Optimal strategies for seeking information: models for statistics, choice reaction times, and human information processing. *J.math.Psychol.*, 1965, 2, 312-329
- EGAN, J.P.: Recognition memory and the operating characteristic. Indiana Univer., Hearing and Communication Lab., Tech.Note AFCRC-TN-58-51, 1958
- EIMAS, P.D., ZEAMAN, D.: Response speed changes in an Estes' paired-associate "miniature" experiment. *J.verb.Learn. verb.Behav.*, 1963, 1, 384-388
- EIMAS, P.D.: Subjective reinforcement in the paired-associate learning of retarded and normal children. *Canad.J.Psychol./Rev.Canad.Psychol.*, 1964, 18, 183-196
- EKSTRAND, B.R., WALLACE, W.P., UNDERWOOD, B.J.: A frequency theory of verbal discrimination learning. *Psychol.Rev.*, 1966, 73, 566-578
- ERICKSON, J.R.: Effects of punishment for errors on discrimination learning by humans. *J.exp.Psychol.*, 1970, 83, 112-119

- ERICKSON, J.R., ZAJKOWSKI, M.M.: Learning several concept identification problems concurrently: A test of the sampling-with-replacement assumption. *J.exp.Psychol.*, 1967, 74, 212-218
- ERICKSON, J.R., ZAJKOWSKI, M.M., EHMANN, E.D.: All-or-none assumptions in concept identification: Analysis of latency data. *J.exp.Psychol.*, 1966, 72, 690-697
- ESTES, W.K.: Toward a statistical theory of learning. *Psychol.Rev.*, 1950, 57, 94-107
- The statistical approach to learning theory. In: KOCH, S. (Ed.): *Psychology: a study of a science*. Vol. II. New York: McGraw-Hill, 1959, 380-491
 - Reinforcement in human learning. In: TAPP, J.T., 1969, 63-94
 - Reward in human learning: Theoretical issues and strategic choice points. In: GLASER, R., 1971, 16-36
- FAIMAGNE, J.C., THEIOS, J.: On attention and memory in reaction time experiments. *Acta Psychol.*, 1969, 30, 316-323
- FERGUSON, E.L., BUSS, A.H.: Supplementary report: Acquisition, extinction, and counterconditioning with different verbal reinforcement combinations. *J.exp.Psychol.*, 1959, 58, 94-95
- FITTS, P.M.: Cognitive aspects of information processing: III. Set for speed versus accuracy. *J.exp.Psychol.*, 1966, 71, 849-857
- FOPPA, K.: *Lernen, Gedächtnis, Verhalten*. Köln: Kiepenheuer und Witsch, 1965
- GAGNÉ, R.M.: *Die Bedingungen des menschlichen Lernens*. Hannover: Schroedel, 1969
- GLASER, R. (Ed.): *The nature of reinforcement*. New York: Academic Press, 1971
- GREENO, J.G., STEINER, T.E.: Markovian processes with identifiable states: General considerations and applications to all-or-none learning. *Psychometrika*, 1964, 29, 309-333
- GREENWALD, A.G.: Nuttin's neglected critique of the law of effect. *Psychol.Bull.*, 1966, 65, 199-205

- GREENWALD, A.G.: Difficulty of associative performance following training with negative instances. A note on punishment effects. *J.educ.Psychol.*, 1970, 61, 255-259
- GUTHRIE, J.T.: Feedback and sentence learning. *J.verb.Learn.verb.Behav.*, 1971, 10, 23-28
- HALE, D.: The relation of correct and error responses in a serial choice reaction time task. *Psychon.Sci.*, 1968, 13, 299-300
- HALL, R.F.: The effect of number of response alternatives on response frequency and latency in paired-associate learning. *Brit.J.math.statist.Psychol.*, 1969, 22, 115-130
- HALL, R.F., WENDEROTH, P.M.: Instruction induced strategies in paired-associate learning. *J.verb.Learn.verb.Behav.*, 1970, 9, 308-310
- Effects of number of responses and recall strategies on parameter values of a paired-associate learning model. *J.verb.Learn.verb.Behav.*, 1972, 11, 29-37
- HARLEY, W.F., Jr.: The effect of monetary incentive in paired-associate learning using a differential method. *Psychon.Sci.*, 1965 a, 2, 377-378
- The effect of monetary incentive in paired-associate learning using an absolute method. *Psychon.Sci.*, 1965 b, 3, 141-142
- HARLOW, H.F.: The formation of learning sets. *Psychol.Rev.*, 1949, 56, 51-65
- HAYS, W.L.: Statistics. New York: Holt, Rinehart and Winston, 1963
- HEIDE, M.: Relative Effizienz von positiver und negativer verbaler Rückmeldung im Lernexperiment. Unveröffentl. Diplomarbeit, Fachbereich Psychologie der Univ.Marburg, 1974
- HERRMANN, TH., STAFF, K.H.: Positive und negative Rückmeldung bei Alternativdiskriminationen: Rückmeldungsproportion und Lernstrategie. *Psychol.Forsch.*, 1973, 36, 117-143
- HILGARD, E.R., BOWER, G.H.: Theorien des Lernens I u. II. Stuttgart: Klett, 1973 (Amerikanische Originalausgabe, 3. Aufl.: Theories of learning. 1966)

- HILLIX, W.A., MARX, M.H.: Response strengthening by information and effect in human learning. *J.exp.Psychol.*, 1960, 60, 97-102
- HOLLANDER, M., WOLFE, D.A.: Nonparametric statistical methods. New York: Wiley, 1973
- HOLMAN, E.W.: The relative effects of reward and nonreward in simple learning. *J.math.Psychol.*, 1969, 6, 497-513
- HULL, C.L.: Principles of behavior. New York: Appleton-Century-Crofts, 1943
- HUMPHREYS, M.S.: Recall and repetition in a modified right-wrong paradigm. *J.verb.Learn.verb.Behav.*, 1972, 11, 573-582
- IBES, K., KELLER, R.: Einfluß der Rückmeldung beim Paarassoziationslernen auf Wiederholungshäufigkeiten und Latenzzeiten. Teil I. Unveröffentl. Semesterarbeit, Fachbereich Psychologie der Univ. Marburg, 1973
- JONES, A.: The relative effectiveness of positive and negative verbal reinforcers. *J.exp.Psychol.*, 1961, 62, 368-371
- KELLER, L., COLE, M., BURKE, C.J., ESTES, W.K.: Reward and information values of trial outcomes in paired-associate learning. *Psychol. Monogr.*, 1965, 79 (Whole No. 605)
- KELLER, L., THOMSON, W.J., TWEEDY, J.R., ATKINSON, R.C.: Effects of reinforcement intervals in paired-associate learning. *J.exp.Psychol.*, 1967, 73, 268-277
- KINTZ, B.L., PAPPAS, J.P.: A scaling evaluation of the effectiveness of verbal reinforcement. *J.Psychol.*, 1965, 59, 359-365
- KREYSZIG, E.: Statistische Methoden und ihre Anwendungen. Göttingen: Vandenhoeck u. Ruprecht, 1965
- LaBERGE, D.L.: A recruitment theory of simple behavior. *Psychometrika*, 1962, 27, 375-396
- LAHEY, B.B., VERPLANCK, W.S.: Contiguity and reinforcement factors in multiple-choice verbal learning: Parametric influences. *Psychon.Sci.*, 1970, 19, 93-95

- LAMING, D.R.J.: A new interpretation of the relation between choice-reaction time and the number of equiprobable alternatives. *Brit.J.math.statist.Psychol.*, 1966, 19, 139-149
- Information theory of choice-reaction times. New York: Academic Press, 1968
- LEVINE, M.: Latency-choice discrepancy in concept learning. *J.exp.Psychol.*, 1969, 82, 1-3
- LEVINE, M., LEITENBERG, H., RICHTER, M.: The blank trials law: The equivalence of positive reinforcement and non-reinforcement. *Psychol.Rev.*, 1964, 71, 94-103
- LIENERT, G.A.: Verteilungsfreie Methoden in der Biostatistik. Band I. Meisenheim am Glan: Hain, 1973
- LIU, I.: Effects of "right" and "wrong" in a Thorndikian experiment. *J.exp.Psychol.*, 1972, 95, 365-369
- LONGSTRETH, L.E.: Tests of the law of effect using open and closed tasks. *J.exp.Psychol.*, 1970, 84, 53-57
- Relationship between response learning and recall of feedback in tests of the law of effect. *J.exp.Psychol.*, 1971, 90, 149-151
- LONGSTRETH, L.E., BARGE, G.: Recall of feedback as it affects verbal learning in tests of the law of effect. *Proc.Ann. Convention Amer.Psychol.Ass.*, 1970, 5, 67-68
- LOVELACE, E.A.: Knowledge of correct and incorrect items in verbal discrimination. *Psychon.Sci.*, 1966, 6, 363-364
- MARX, M.H.: Increased probability of error repetition as a function of number of successive prior repetitions. *Percept.Mot.Skills*, 1971, 32, 544-546
- MARX, M.H., MARX, K.: Observation vs. performance in learning over the fourth to sixth grades. *Psychon.Sci.*, 1970, 21, 199-200
- MARX, M.H., WITTER, D.W.: Repetition of correct responses and errors as a function of performance with reward or information. *J.exp.Psychol.*, 1972, 92, 53-58
- MCGILL, W.J.: Stochastic latency mechanisms. In: LUCE, R.D., BUSH, R.R., GALANTER, E. (Eds.): *Handbook of mathematical psychology*. Vol. I. New York: Wiley, 1963, 309-360

- McGILL, W.J., GIBBON, J.: The general-gamma distribution and reaction times. *J.math.Psychol.*, 1965, 2, 1-18
- MEDIN, D.L.: Partial information and choice behavior in differential reward magnitude learning *Psychon.Sci.*, 1972, 27, 73-76
- MEYER, W.J., OFFENBACH, S.I.: Effectiveness of reward and punishment as a function of task complexity. *J.comp. physiol.Psychol.*, 1962, 55, 532-534
- MILLWARD, R.: Latency in a modified paired-associate learning experiment. *J.verb.Learn.verb.Behav.*, 1964 a, 3, 309-316
- An all-or-none model for noncorrection routines with elimination of incorrect responses. *J.math.Psychol.*, 1964 b, 1, 392-404
- MONTAGUE, W.E., KIESS, H.O.: The associability of CVC pairs. *J.exp.Psychol.Monogr.Suppl.*, 1968, 78 (2, Pt. 2)
- MOSBERG, L.: Response elimination in noncorrection paired-associate learning. *J.exp.Psychol.*, 1970, 83, 94-100
- MUELLER, C.G.: Theoretical relationships among some measures of conditioning. *Proc.natl.Acad.Sci.*, 1950, 36, 123-130
- MUELLER, D.J., GUMINA, J.M.: Methodological studies involving reward and punishment in human learning: I. Effects of design and instructions. *Percept.Mot.Skills*, 1972, 34, 63-74
- MURDOCK, B.B., Jr.: Short-term memory for associations. In: NORMAN, D.A., 1970, 285-304
- MURDOCK, B.B., Jr., DUFTY, P.O.: Strength theory and recognition memory. *J.exp.Psychol.*, 1972, 94, 284-290
- MYERS, J.L., ATKINSON, R.C.: Choice behavior and reward structure. *J.math.Psychol.*, 1964, 1, 170-203
- MYERS, J.L., GAMBINO, B., JONES, M.R.: Response speeds in probability learning. *J.math.Psychol.*, 1967, 4, 473-488
- NAHINSKY, I.D.: A duoprocess model for paired-associate learning. *Psychol.Rep.*, 1964, 14, 467-471
- All-or-none elimination of errors in paired-associate learning. *J.verb.Learn.verb.Behav.*, 1966, 5, 397-401
 - Statistics and moments-parameter estimates for a duoprocess paired-associate learning model. *J.math.Psychol.*, 1967, 4, 140-150

- NAHINSKY, I.D., MUELLER, J.H.: A test of two all-or-none models for paired-associate learning. *J.verb.Learn.verb. Behav.*, 1968, 7, 464-473
- NICKERSON, R.S.: "Same"-"different" response times: A model and a preliminary test. *Acta Psychol.*, 1969, 30, 257-275
- NORMAN, D.A. (Ed.): *Models of human memory*. New York: Academic Press, 1970
- NORMAN, D.A., WICKELGREN, W.A.: Strength theory of decision rules and latency in retrieval from short-term memory. *J.math.Psychol.*, 1969, 6, 192-208
- NUTTIN, J.: *Tâche, réussite et échec*. Louvain, Belgium: Publications Univ. Louvain, 1953
- NUTTIN, J., GREENWALD, A.G.: *Reward and punishment in human learning*. New York: Academic Press, 1968
- PARREREN, C.F. van: *Lernprozeß und Lernerfolg*. Braunschweig: Westermann, 1966
- PETERSON, L.R.: Paired-associate latencies after the last error. *Psychon.Sci.*, 1965, 2, 167-168
- PIKE, A.R.: Latency and relative frequency of response in psychophysical discrimination. *Brit.J.math.statist.Psychol.*, 1968, 21, 161-182
- The latencies of correct and incorrect responses in discrimination and detection tasks: Their interpretation in terms of a model based on simple counting. *Percept. and Psychophys.*, 1971, 9, 455-460
- POLLAY, R.W.: A model of decision times in difficult decision situations. *Psychol.Rev.*, 1970, 77, 274-281
- POSTMAN, L.: Rewards and punishments in human learning. In: POSTMAN, L. (Ed.): *Psychology in the making*. New York: Knopf, 1962, 331-401
- Reply to Greenwald. *Psychol.Bull.*, 1966, 65, 383-388
- POSTMAN, L., ADAMS, P.A.: Performance variables in the experimental analysis of the law of effect. *Amer.J.Psychol.*, 1954, 67, 612-631
- "Isolation" and the law of effect. *Amer.J.Psychol.*, 1955, 68, 96-105

- REITMAN, W.: What does it take to remember? In: NORMAN, D.A., 1970, 469-509
- RESTLE, F.: The selection of strategies in cue learning. *Psychol.Rev.*, 1962, 69, 329-343
- RESTLE, F., GREENO, J.G.: Introduction to mathematical psychology. Reading, Mass.: Addison-Wesley, 1970
- RIMM, D., ROESCH, R., PERRY, R., PEEBLES, CH.: Effects of blank versus noninformative feedback and "right" and "wrong" on response repetition in paired-associate learning. *J.exp.Psychol.*, 1971, 88, 26-30
- SACHS, L.: Statistische Auswertungsmethoden. Berlin: Springer, 1969
- SCHEIRER, C.J.: Effect of cueing, modality, and effective contiguous time on response latency in short-term memory. *J.exp.Psychol.*, 1971, 88, 429-432
- SCHOUTEN, J.F., BEKKER, J.A.M.: Reaction time and accuracy. *Acta Psychol.*, 1967, 27, 143-153
- SCHWARZE, H.: Wirksamkeit und Wirkungsweise von Lob und Tadel bei einfachen Lernaufgaben. Unveröffentl. Diss., Fachbereich Psychologie der Univ. Marburg, 1973
- SHANNON, C.E., WEAVER, W.: The mathematical theory of communication. Urbana: Univ. of Illinois Press, 1949
- SHIFFRIN, R.M.: Memory search. In: NORMAN, D.A., 1970, 375-447
- SHIFFRIN, R.M., ATKINSON, R.C.: Storage and retrieval processes in long-term memory. *Psychol.Rev.*, 1969, 76, 179-193
- SIXTL, F.: Meßmethoden der Psychologie. Weinheim: Beltz, 1967
- SMITH, E.E.: Choice reaction time: An analysis of the major theoretical positions. *Psychol.Bull.*, 1968, 69, 77-110
- SNODGRASS, J.G., LUCE, R.D., GALANTER, E.: Some experiments on simple and choice reaction time. *J.exp.Psychol.*, 1967, 75, 1-17
- SPENCE, J.T.: Verbal discrimination performance under different verbal reinforcement combinations. *J. exp.Psychol.*, 1964 a, 67, 195-197

- SPENCE, J.T.: Performance on a four-alternative task under different reinforcement combinations. *Psychon.Sci.*, 1964 b, 1, 241-242
- Effects of verbal reinforcement combination and instructional condition on performance of a problem-solving task. *J.Pers.soc.Psychol.*, 1966 a, 3, 163-170
 - Verbal-discrimination performance as a function of instructions and verbal reinforcement combination in normal and retarded children. *Child Developm.*, 1966 b, 37, 269-281
 - The effects of verbal reinforcement combination on the performance of a four-alternative discrimination task. *J.verb.Learn.verb.Behav.*, 1966 c, 5, 421-428
 - Verbal reinforcement combinations and concept-identification learning: The role of nonreinforcement. *J.exp. Psychol.*, 1970, 85, 321-329
 - Effects of blank versus noninformative feedback and "right" and "wrong" on response repetition in paired-associate learning: A reanalysis and reinterpretation. *J.exp.Psychol.*, 1972, 94, 146-148
- SPENCE, J.T., LAIR, C.V., GOODSTEIN, L.D.: Effects of different feedback conditions on verbal discrimination learning in schizophrenic and nonpsychiatric subjects. *J.verb.Learn.verb.Behav.*, 1963, 2, 339-345
- SPENCE, J.T., LAIR, C.V.: The effect of different verbal reinforcement combinations on the verbal discrimination performance of schizophrenics. *J.Pers.soc.Psychol.*, 1965, 1, 245-249
- STEPHENS, J.M.: The influence of punishment on learning. *J. exp.Psychol.*, 1934 a, 17, 536-555
- Further notes on punishment and reward. *J.genet.Psychol.*, 1934 b, 44, 464-472
- STERNBERG, S.: High-speed scanning in human memory. *Science*, 1966, 153, 652-654
- The discovery of processing stages: Extensions of Donder's method. *Acta Psychol.*, 1969 a, 30, 276-315

- STERNBERG, S.: Memory-scanning: Mental processes revealed by reaction-time experiments. Amer.Scientist, 1969 b, 57, 421-457
- STONE, M.: Models for choice-reaction time. Psychometrika, 1960, 25, 251-260
- SUPPES, P., GROEN, G., SCHLAG-REY, M.: A model for response latency in paired-associate learning. J.math.Psychol., 1966, 3, 99-128
- TAPP, J.T. (Ed.): Reinforcement and behavior. New York: Academic Press, 1969
- THEIOS, J.: Predictions of paired-associate latencies after the last error by an all-or-none learning model. Psychon. Sci., 1965, 2, 311-312
- Mathematical models for aversive conditioning. In: BRUSH, F.R. (Ed.): Aversive conditioning and learning. New York: Academic Press, 1971, 297-345
- THOMAS, E.A.C.: Sufficient conditions for monotone hazard rate: An application to latency-probability curves. J. math.Psychol., 1971, 8, 303-332
- THORNDIKE, E.L.: Animal intelligence: experimental studies. New York: Macmillan, 1911
- The fundamentals of learning. New York: Teachers Coll., Columbia Univ., 1932
- TILTON, J.W.: The effect of "right" and "wrong" upon the learning of nonsense syllables in multiple choice arrangement. J.educ.Psychol., 1939, 30, 95-115
- TOLMAN, E.C.: Purposive behavior in animals and men. New York: Appleton-Century-Crofts, 1932
- TRABASSO, T., BOWER, G.H.: Attention in learning. New York: Wiley, 1968
- WAUGH, N.C.: Associative symmetry and recall latencies: A distinction between learning and performance. Acta Psychol., 1970, 33, 326-337
- WEARING, A.J.: On the trace strength of responses varying in correctness. Psychon.Sci., 1970, 21, 226-227
- On the Adams-Bray retrieval model. J.exp.Psychol., 1971, 89, 96-101

- WICKELGREN, W.A.: Multitrace strength theory. In: NORMAN, D.A., 1970, 65-102
- WILLIAMS, J.P.: A test of the all-or-none hypothesis for verbal learning. J.exp.Psychol., 1962, 64, 158-165
- WINER, B.J.: Statistical principles in experimental design. London: McGraw-Hill, 1970
- WITTER, D.W., MUELLER, J.H., MARX, M.H.: Correction procedures in observational learning. Psychon.Sci., 1971, 22, 94-95
- WOLLEN, K.A.: Effects of set to learn A-B or B-A upon A-B and B-A tests. J.exp.Psychol., 1970, 86, 186-189
- WOODWORTH, R.S., SCHLOSBERG, H.: Experimental psychology. New York: Holt, Rinehart and Winston, 1954
- YELLOTT, J.I., Jr.: Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. J.math. Psychol., 1971, 8, 159-199
- ZIRKLE, G.A.: Success and failure in serial learning. II. Isolation and the Thorndike effect. J.exp.Psychol., 1946, 36, 302-315

11. Anhang

11.1 Lernlisten

Experimente I und II: Die Lernlisten sind für die Bedingung $k=4$ angegeben; für $k=2$ ist 7 durch 3 und 9 durch 5 zu ersetzen. Die aufgeführten Permutationen der Reizwortreihenfolgen wurden in den ungeradzahligten Lerndurchgängen benutzt; weitere Permutationen findet man bei IBES u. KELLER (1973).

<u>Liste a</u>	<u>Liste A</u>	<u>Liste b</u>	<u>Liste B</u>
ZEIT 7 r	(Forts.)	LEIB 5 f	(Forts.)
HAUS 5 f	HALT 3 r	SPUR 7 r	KERL 9 f
SOHN 9 f	RUHM 7 b	KERN 3 b	SIEG 7 b
BANK 3 r	NEST 9 f	WAND 9 f	NEID 9 f
HEFT 5 f	GLAS 3 r	TIER 3 r	MORD 3 f
GRAD 3 b	BEIN 7 b	GELD 5 b	NETZ 3 b
FORM 7 r	GOLD 5 f	LAST 7 f	HALS 9 r
MARK 9 r	WITZ 9 r	SITZ 9 r	MEHL 5 f
BILD 9 b	REIS 7 r	HEIL 3 f	QUAL 3 r
VOLK 7 f	ZORN 9 b	TUCH 7 r	ZEUG 7 r
HERZ 5 b	KURS 3 r	WERT 5 r	BAUM 9 b
TEIL 3 f	RING 5 b	GLUT 9 b	VERS 5 b
BROT 5 r	PARK 7 r	SALZ 5 f	DUFT 7 f
WERK 7 b	ROHR 9 f	ROCK 7 b	RAUB 9 r
SATZ 9 f	HELD 5 r	DANK 3 r	WINK 3 f
LAND 3 r	WIND 3 f	ZELT 9 f	GRAS 5 r
	FACH 7 f		MAGD 3 b
	POST 9 b		BURG 7 f
	JAGD 3 b		KORN 7 r
	BERG 5 r		MAUL 5 f
	HAND 7 f		LOCH 3 r
<u>Liste A</u>		<u>Liste B</u>	
WORT 3 f		STIL 9 r	
KALK 9 r		FLUT 7 f	
ZIEL 5 b		KAUF 3 r	
GIFT 7 f		FELS 5 b	
RAUM 5 r		REIZ 5 f	
BUCH 7 b		HUND 9 r	
DORF 5 b		GAST 7 b	
BAHN 9 r		WILD 3 f	
WELT 9 f		RAND 7 r	
GRAF 3 r		TEXT 5 b	
BIER 7 r		STAB 3 r	
JAHR 9 b		HEIM 9 b	
REST 3 f		LIST 7 b	
PLAN 3 b		BUND 5 r	
MOND 5 f		ZELT 9 f	
LUST 7 f		DING 3 r	
HOLZ 5 r		TURM 7 f	
LEID 3 b		LIED 5 f	
GANG 9 r		DACH 9 b	
BLUT 5 f		CHOR 5 r	

Experimente III, IV und V: Die Lernlisten sind für die beob-
4-Bedingung angegeben; für k=2 ist 7 durch 5 und 9 durch 3
zu ersetzen. Für die "aktiven" Gruppen entfällt selbstver-
ständlich die Vorgabe der Antwortzahlen.

<u>Liste a'</u>		<u>Liste A'</u>		<u>Liste B'</u>	
Lernliste	Prüf- liste	Lernliste	Prüf- liste	Lernliste	Prüf- liste
WERT 5 f	WERT	WIND 9 r	WIND	GLAS 9 b	GLAS
FELD 3 r	FELD	BURG 5 b	BURG	BERG 7 r	BERG
LAST 9 b	MEER	GARN 5 f	HERZ	LEID 5 r	ZAHN
WITZ 9 r	LEIB	LOCH 3 f	ZAHL	HOLZ 3 f	TRUG
RAUM 7 f	NEST	KNIE 9 r	WERK	KRUG 7 f	BALL
BOOT 5 b	LAST	POST 3 b	KOPF	SATZ 9 b	BLUT
TEXT 3 b	TEXT	TUCH 7 f	SAAT	TOFF 7 r	HOLZ
SPUR 3 f	RAUM	ZAHL 5 b	ZORN	ZEIT 9 b	ZEIT
SOHN 7 b	SOHN	BROT 9 r	KORB	BALL 5 b	SATZ
HALT 5 r	WITZ	FELL 3 r	KNIE	MEHL 3 b	GELD
NEST 7 r	SPUR	GRAS 7 b	LIST	ZAHN 5 f	DARM
LEIB 9 f	BOOT	HERZ 9 f	VIEH	BLUT 7 f	KRUG
BUCH 3 b	HALT	KORB 5 r	LOCH	FORM 3 r	ZEUG
MEER 5 b	BUCH	LIST 9 b	BROT	HEFT 3 b	FORM
SIEB 9 r	SIEB	SAAT 3 r	GRAS	DARM 5 b	MEHL
TEIL 7 f	TEIL	VIEH 3 f	TUCH	RING 9 f	WALD
		WERK 7 f	BAHN	GELD 7 r	LEID
		ZORN 5 b	GARN	ZEUG 5 r	HEFT
		BAHN 5 b	FELL	WALD 3 f	TOFF
		KOFF 7 r	BILD	TRUG 9 r	RING
		SAUM 3 b	SAUM	ROST 7 b	KAMM
		LUFT 7 b	KIND	KAMM 3 b	PFAD
		TAKT 9 f	POST	SIEG 9 f	WEIN
		ZIEL 5 f	DACH	VOLK 5 r	TIER
		MUND 9 r	LUFT	HAUT 7 f	ROST
		QUAL 3 r	STAB	PFAD 9 r	GRAB
		BILD 3 b	MUND	ZOLL 3 r	ZOLL
		FUND 9 b	HAUS	WEIN 7 b	LUST
		HAAR 5 r	ROCK	HAST 3 f	MOOR
		ROCK 7 f	TAKT	TIER 9 f	SIEG
		ZELT 7 r	FUND	LAND 5 b	HAST
		TURM 9 b	ZELT	KOHL 5 f	VOLK
		STAB 5 f	VERB	NETZ 7 b	LAND
		VERB 3 f	QUAL	GRAB 9 r	BANK
		DACH 7 r	MARK	PLAN 3 r	KOHL
		KIND 9 f	TURM	MOOR 9 b	HAUT
		MARK 7 b	ZIEL	BANK 7 r	NETZ
		HAUS 5 r	HAAR	LUST 5 f	PLAN
		GLUT 3 b	GLUT	SACK 3 f	SACK
		HEMD 7 f	HEMD	GROG 5 b	GROG

1.2 Instruktionen

Experiment I

Hauptinstruktion: s.S. 77-80

R-recall-Instruktion:

"Im folgenden Prüfdurchgang sollst Du nicht wie bisher die dem jeweiligen Reizwort entsprechende, möglichst richtige Zahl nennen, sondern die Zahl wiederholen, die in den Lerndurchgängen auf den Karten stand, unabhängig davon, ob die Rückmeldung richtig, falsch oder "—" lautete."

O-recall-Instruktion:

"Im folgenden Prüfdurchgang sollst Du nicht wie bisher die dem jeweiligen Reizwort entsprechende, möglichst richtige Zahl nennen, sondern nur die Rückmeldung wiederholen (richtig, falsch oder "—"), die bei dem jeweiligen Reizwort in den Lerndurchgängen auf den Karten stand. Du sollst also keine Zahl mehr nennen!"

R-remind-Instruktion:

"Die L e r n durchgänge bleiben unverändert!

Der P r ü f durchgang dagegen wird in folgender Weise abgeändert:

Bei der Darbietung einiger, zufällig ausgewählter Reizwörter wirst Du zusätzlich an die Antwort aus den Lerndurchgängen - als Gedächtnisstütze - erinnert werden. Die Antwort (Zahlalternative) steht an der gleichen Stelle wie im Lerndurchgang und ist eingeklammert.

Beispiel für die unterschiedliche Darbietung im Prüfdurchgang: Wenn z.B. in den Lerndurchgängen auf einer Karte die Folge

"T E I G 3 falsch"

gestanden hatte, wird im Prüfdurchgang

"T E I G (3)"

dargeboten werden.

Es bedeutet also nicht, daß die erinnerten Antworten unbedingt richtig sein müssen, sondern lediglich, daß eben diese Zahl im Lerndurchgang an dieser Stelle stand.

Wie bisher sollst Du im Prüfdurchgang bei Deinen Antworten möglichst wenig Fehler machen!

Hast Du noch Fragen? - Wiederhole bitte kurz, wie der modifizierte Prüfdurchgang ablaufen wird!"

O-remind-Instruktion:

"Die L e r n durchgänge bleiben unverändert!

Der P r ü f durchgang dagegen wird in folgender Weise abgeändert:

Bei der Darbietung einiger, zufällig ausgewählter Reizwörter wirst Du zusätzlich an die Rückmeldung aus den Lerndurchgängen - als Gedächtnisstütze - erinnert werden. Die Rückmeldung steht an der gleichen Stelle wie im Lerndurchgang und ist eingeklammert.

Beispiel für die unterschiedliche Darbietung im Prüfdurchgang: Wenn z.B. in den Lerndurchgängen auf einer Karte die Folge

"T E I G 3 falsch"

gestanden hatte, wird im Prüfdurchgang

"T E I G (falsch)"

dargeboten werden.

Wie bisher sollst Du im Prüfdurchgang bei Deinen Antworten möglichst wenig Fehler machen!

Hast Du noch Fragen? - Wiederhole bitte kurz, wie der modifizierte Prüfdurchgang ablaufen wird!"

Experiment II

"Symmetriebetonende" Instruktion:

"Da Du schon an einem ähnlichen Lernversuch teilgenommen hast, kann die Instruktion relativ kurz gehalten werden. Das Lernmaterial wird wieder auf Kärtchen (s.Beispiele) dargeboten werden, die Du selber im Rhythmus eines Metronoms lesen und wechseln wirst.

Den einzelnen Reizworten ist nach Zufall je eine der 4 Zahlalternativen (3,5,7 und 9) als Antwort zugeordnet worden. Auf jedem Kärtchen steht nach Reizwort und Zahlalternative eine "Rückmeldung", aus der Du entnehmen kannst, ob die angegebene Zahlalternative die "richtige" oder "falsche" Ant-

wort auf das Reizwort ist.

Die Rückmeldung "—" (lies: blänk) gibt keine Auskunft darüber, ob die Zahlalternative richtig oder falsch ist.

Nach einer Reihe von Lerndurchgängen werden Dir im Prüfungsdurchgang die Reizworte einzeln dargeboten. Versuche, auf die Reizworte möglichst richtig zu antworten, d.h. Zahlalternativen, denen die Rückmeldung "richtig" gefolgt war, zu wiederholen, bzw. Zahlalternativen, denen die Rückmeldung "falsch" gefolgt war, nicht mehr zu wiederholen.

(Während des Prüfungsdurchgangs selbst bekommst Du keine Rückmeldung; es werden nur Deine Antworten registriert.)

Du erhältst im ganzen Versuch drei verschiedene Listen. Für jede Liste hast Du vier Lerndurchgänge zur Verfügung; die Reihenfolge der Kärtchen ist dabei von Durchgang zu Durchgang verschieden.

Das Hantieren mit den Kärtchen im Rhythmus des Metronoms wurde eingeführt, um eine aufwendige Projektionsapparatur zu vermeiden. Damit dieses Hantieren nicht vom eigentlichen Lernen abhält, soll es vorher noch einmal mit dem Benennen von Farbpunkten geübt werden. - Die Befolgung des Metronomtakts gilt nicht für den Prüfungsdurchgang!

(Das Tonbandgerät dient wieder zur Kontrolle meines Protokolls.)

Sinn der Wiederholung des Experiments:

Das vorhergehende Experiment hat ergeben, daß es bei vier Zahlalternativen anscheinend ziemlich schwierig ist, "richtige" Alternativen zu wiederholen bzw. "falsche" nicht mehr zu wiederholen. Deshalb wird nun ein vierter Lerndurchgang eingeführt. Es nehmen nur Versuchspersonen teil, die schon an die ungewöhnliche Darbietung des Lernmaterials gewöhnt sind.

Hast Du noch Fragen zum Versuchsablauf? Fragen zum theoretischen Hintergrund kann ich Dir erst nach Beendigung des Versuchs beantworten."

Experimente III und IV

Hauptinstruktion:

"In diesem Lernversuch besteht Deine Aufgabe darin, zu einer Liste von einsilbigen deutschen Hauptworten (den sog. Reizworten) die zugehörige Antwort zu erlernen. Als mögliche Antwortalternativen kommen die Zahlen 3, 5, 7, 9 in Frage. Sie sind den Reizworten mit Hilfe einer Lotterietrommel zufällig zugeordnet worden. Es gibt also in der Liste mehrere Reizworte, denen die Zahl 3 zugeordnet worden ist, andere wiederum, denen 5 zugeordnet worden ist, usw.

Du erhältst einen Lerndurchgang und einen Prüfdurchgang. Im Lerndurchgang wird Dir die Liste Wort für Wort auf dem Projektionsschirm nacheinander dargeboten werden. Sage zu jedem Reizwort eine der möglichen Antworten (3, 5, 7, 9). Da Du die richtige Zuordnung nicht kennst, mußt Du zunächst raten. Versuche dabei, keine der Zahlen irgendwie zu bevorzugen. Nach jeder Antwort erhältst Du von mir eine Rückmeldung ("richtig" oder "falsch"), ob Du die Zuordnung getroffen hast oder nicht (was Dir, wie gesagt, nur per Zufall gelingen kann).

Beim Prüfdurchgang werden Dir die Reizworte wieder dargeboten, und zwar in veränderter Reihenfolge. Antworte wieder mit einer Zahl. Bemühe Dich dabei, die Zahlen, denen im Lerndurchgang ein "richtig" gefolgt war, zu wiederholen, und die Zahlen, denen ein "falsch" gefolgt war, zu vermeiden. Im letzteren Falle nenne dann irgendeine andere der Zahlalternativen.

Die Punktzahl, die Deinen Lernerfolg angibt, setzt sich also sowohl aus "richtig"-Zahl-Wiederholungen als auch aus "falsch"-Zahl-Vermeidungen zusammen.

Beispiel

<u>Lerndurchgang</u>			<u>Prüfdurchgang</u>		
<u>Reizwort</u>	<u>Zahl</u>	<u>Rückmeldung</u>	<u>Reizwort</u>	<u>Zahl</u>	
FLUT	5	falsch	FLUT	3	(Lernerfolg!)
WILD	7	richtig	WILD	5	(<u>kein</u> Lernerfolg!)
RAUB	9	falsch	RAUB	9	(<u>kein</u> Lernerfolg!)
STIL	3	richtig	STIL	3	(Lernerfolg!)

Zusätzlich zu den Rückmeldungen "richtig" und "falsch" im Lerndurchgang gibt es noch eine dritte "Rückmeldung", nämlich "blank" (gesprochen: blänk), z.B. "HALS 9 blank".

Dieser Ausdruck ist aus der englischen Fachliteratur übernommen und besagt, daß an dieser Stelle der Versuchsleiter der Versuchsperson keine Rückmeldung darüber gibt, ob ihre Antwort "richtig" oder "falsch" war.

(Die Einführung der "blank"-Rückmeldung bezieht sich auf alltägliche Lernsituationen, in denen ebenfalls eine unmittelbare Rückmeldung ausbleibt.)

Wenn Du also im Lerndurchgang bei einigen Reizworten und den von Dir genannten Zahlen die Rückmeldung "blank" erhältst, kannst Du demnach nicht wissen, ob Du diese Zahlen im Prüf-durchgang vermeiden oder wiederholen sollst. Es bleibt Dir in diesen Fällen nichts anderes übrig, als im Prüfdurchgang einfach wieder eine der Zahlen 3, 5, 7, 9 auszuwählen. - Natürlich ist aufgrund einer "blank"-Rückmeldung kein Beitrag zum Lernerfolg möglich.

Im Prüfdurchgang gibt es auf Deine Antworten keine Rückmeldungen mehr. (Nach Abschluß des Versuchs kann ich Dir aber die Ergebnisse zeigen.)

Der Versuch läuft also wie folgt ab:

Lerndurchgang:

1. Auf dem Projektionsschirm erscheint kurzzeitig ein Reizwort.
2. Antworte darauf mit einer der Zahlen 3, 5, 7, 9, die Du dabei einigermaßen zufällig auswählen sollst.
3. Du erhältst daraufhin von mir eine Rückmeldung: "richtig" oder "falsch" oder "blank".
4. Das nächste Reizwort erscheint usw., bis die Liste (40 Worte) beendet ist.

Prüfdurchgang:

5. Die Reizworte erscheinen in veränderter Reihenfolge wieder auf dem Schirm.
6. Versuche so zu antworten, daß Du nach Möglichkeit "richtig"-Zahlen wiederholst bzw. "falsch"-Zahlen vermeidest. Beides stellt einen Lernerfolg dar.

7. Gib auf jeden Fall eine Antwort (3, 5, 7, 9), auch wenn Du meinst, Dich bei einigen Reizworten an überhaupt nichts mehr erinnern zu können.
8. Sprich, bitte, laut, da das Mikrophon, durch das der Projektionsapparat weitergeschaltet wird, erst ab einer gewissen Lautstärke anspricht.

Hast Du noch Fragen zum Versuchsablauf? - Fragen, die sich auf den theoretischen Hintergrund des Experiments beziehen, kann ich Dir erst nach Abschluß des Versuchs beantworten."

Experiment V

Zunächst wurde die Hauptinstruktion aus Experiment III in der k=2-Version gegeben (die letzten drei Zeilen waren abgedeckt); dann folgte die jeweilige Strategieinstruktion.

S-R-f-Instruktion:

"Bei diesem Experiment interessiert besonders die Bedeutung verschiedener Lernstrategien.

So kann man sich entweder

1. alles genau so merken, wie es im Lerndurchgang abläuft (also: Reizwort - Deine Zahl - meine Rückmeldung), oder
2. nur das Reizwort und die richtige Zahl merken (wenn ich bei einer Zahl "falsch" sage, muß man sich also gerade die andere Zahl einprägen).

Du sollst nun im folgenden die erste (unterstrichene) Strategie anwenden!

Sollte "blank" die Rückmeldung sein, kannst Du, wie schon auf S.2¹⁾ gesagt wurde, selbst entscheiden, wie Du vorgehen willst.

Einige Beispiele sollen das verdeutlichen:

<u>Lerndurchgang</u>			<u>Deine Strategie</u>		
<u>Reizwort - Zahl - Rückmeldung</u>			<u>Bitte in der Form merken:</u>		
FLUT	5	richtig	FLUT	5	richtig
WILD	3	falsch	WILD	3	falsch
RAUB	5	blank	(beliebig, wie Du willst)		

Versuche bitte, Dich wirklich an diese Strategie zu halten!

1) S. 2 in der Originalinstruktion entspricht S. 267 in vorliegender Arbeit.

Ordne bitte im Lerndurchgang die Zahlen zufällig und nicht nach einem bestimmten Schema zu (z.B. nicht: 3 immer zu Wörtern mit e und i, 5 zu Wörtern mit o, u oder a; oder auch nicht: 3 immer zu Wörtern mit Anfangsbuchstaben von A bis M, 5 immer zu Wörtern mit Anfangsbuchstaben von N bis Z).

In einem solchen Fall kann ich das Versuchsprotokoll nicht mehr auswerten!

Hast Du noch Fragen zum Versuchsablauf? - Fragen, die sich auf den theoretischen Hintergrund des Experiments beziehen, kann ich Dir erst nach Abschluß des Versuchs beantworten."

S-R-Instruktion:

"Bei diesem Experiment interessiert besonders die Bedeutung verschiedener Lernstrategien.

So kann man sich entweder

1. alles genau so merken, wie es im Lerndurchgang abläuft (also: Reizwort - Deine Zahl - meine Rückmeldung), oder
2. nur das Reizwort und die richtige Zahl merken (wenn ich bei einer Zahl "falsch" sage, mußt Du Dir also gerade die andere Zahl einprägen).

Du sollst nun im folgenden die zweite (unterstrichene) Strategie anwenden!

Sollte "blank" die Rückmeldung sein, kannst Du, wie schon auf S. 2¹⁾ gesagt wurde, selbst entscheiden, wie Du vorgehen willst.

Einige Beispiele sollen das verdeutlichen:

<u>Lerndurchgang</u>			<u>Deine Strategie</u>
<u>Reizwort - Zahl - Rückmeldung</u>			<u>Bitte in der Form merken:</u>
FLUT	5	richtig	FLUT - 5
WILD	3	falsch	WILD - 5 (da 3 ja falsch ist)
RAUB	5	blank	(beliebig, wie Du willst)

Versuche bitte, Dich wirklich an diese Strategie zu halten!

1) S. 2 in der Originalinstruktion entspricht S. 267 in vorliegender Arbeit.

Ordne bitte im Lerndurchgang die Zahlen zufällig und nicht nach einem bestimmten Schema zu (z.B. nicht: 3 immer zu Wörtern mit e und i, 5 zu Wörtern mit o, u oder a; oder auch nicht: 3 immer zu Wörtern mit Anfangsbuchstaben von A bis M, 5 immer zu Wörtern mit Anfangsbuchstaben von N bis Z).

In einem solchen Fall kann ich das Versuchsprotokoll nicht mehr auswerten!

Hast Du noch Fragen zum Versuchsablauf? - Fragen, die sich auf den theoretischen Hintergrund des Experiments beziehen, kann ich Dir erst nach Abschluß des Versuchs beantworten."

Lebenslauf

- 16.12.1939 geboren in Langenbielau/Schlesien
Vater: Ernst Mix (Dipl.-Ing.)
Mutter: Margarete, geb. Zingler
- 1946-1959 Schulzeit. Abschluß durch Abitur (26.2.1959)
am Kaiser-Karls-Gymnasium in Aachen
- 1959 halbjähriges Industriepraktikum
- WS 1959 - Studium der Physik an der TH Aachen (7./8.
SS 1966 Semester an der TU Berlin) - Abschluß mit
Diplom-Hauptprüfung am 10.11.1966
- 1967-1968 Wiss.Angestellter bzw. Assistent am Lehrstuhl
für Experimentalphysik I der TH Aachen (Ar-
beitsgruppe: Sprach- und Musikstatistik)
- SS 1968 - Studium der Psychologie: SS 1968 und WS 1968/
SS 1974 69 an der TH Aachen - ab SS 1969 an der Phi-
lipps-Universität Marburg
- SS 1968 Wiss.Hilfskraft am Lehrstuhl für Psychologie
an der TH Aachen
- WS 1969/70 - Akademischer Tutor am Psychologischen Insti-
SS 1971 tut der Philipps-Universität Marburg (Mitar-
beit an einem Projekt zur Hochschuldidaktik
im Rahmen der Statistikausbildung)
- WS 1971/72 Beginn des Promotionsvorhabens am Fachbereich
04 Psychologie (1.10.1971-30.9.1974 Stipen-
diat nach dem Graduierten-Förderungsgesetz)

